②

AFOSR·TR· o9-ᴜ264

AD-A205 413

Air Force Office of Scientific Research
Life Sciences Directorate

Grant Number 85-0379
9/15/85 - 9/14/88

FINAL REPORT

DTIC
ELECTE
MAR 0 9 1989
S
H

Judith L. Lauter, Ph.D.

Principal Investigator

Speech and Hearing Sciences
University of Arizona
Tucson AZ 85721

Chief, Technical Information Division

2 6 JAN 1989

# REPORT DOCUMENTATION PAGE

| 1a REPORT SECURITY CLASSIFICATION | 1b RESTRICTIVE MARKINGS |
|---|---|
| UNCLASSIFIED | |

| 2a SECURITY CLASSIFICATION AUTHORITY | 3 DISTRIBUTION / AVAILABILITY OF REPORT |
|---|---|
| | Approved for public release; distribution |
| 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE | unlimited. |

| 4 PERFORMING ORGANIZATION REPORT NUMBER(S) | 5. MONITORING ORGANIZATION REPORT NUMBER(S) |
|---|---|
| | AFOSR-TR- 89-0264 |

| 6a. NAME OF PERFORMING ORGANIZATION | 6b OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION |
|---|---|---|
| Univ of Arizona | | Air Force Office of Scientific Research/NL |

| 6c. ADDRESS (City, State, and ZIP Code) | 7b ADDRESS (City, State, and ZIP Code) |
|---|---|
| 1624 E. Helen Street Tucson, AZ 85719 | Building 410 Bolling AFB, DC 20332-6448 |

| 8a. NAME OF FUNDING / SPONSORING ORGANIZATION | 8b. OFFICE SYMBOL (If applicable) | 9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER |
|---|---|---|
| AFOSR | NL | AFOSR-85-0379 |

| 8c. ADDRESS (City, State, and ZIP Code) | 10 SOURCE OF FUNDING NUMBERS | | | |
|---|---|---|---|---|
| | PROGRAM ELEMENT NO. | PROJECT NO. | TASK NO. | WORK UNIT ACCESSION NO |
| Building 410 Bolling AFB, DC 20332-6448 | 61102F | 2313 | A6 | |

**11 TITLE (Include Security Classification)**

Coordinated Noninvasive Studies (CNS) Project

**12. PERSONAL AUTHOR(S)**
Dr. Judith Lauter

| 13a. TYPE OF REPORT | 13b TIME COVERED | 14. DATE OF REPORT (Year, Month, Day) | 15 PAGE COUNT |
|---|---|---|---|
| Final | FROM 15 Sep 85 TO 31 Oct 88 | Nov 88 | 127 |

**16. SUPPLEMENTARY NOTATION**

| 17 | COSATI CODES | | 18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB-GROUP | |
| | | | |
| | | | |

**19 ABSTRACT (Continue on reverse if necessary and identify by block number)**

Research activities during this period include: data collection related to the interface between complex-sound production and perception, specifically, studies on speech acoustics including two experiments on voice-onset-time variability in productions by speakers of several languages, and a series on acoustical characteristics of emotional expression; data collection regarding individual differences in the effect of stimulus characteristic on relative ear advantages; continuing data analysis and new collections documenting individual differences in auditory evoked potentials, with details related to auditory-systems asymmetries preliminary tests regarding the match between behaviroal measures of relative ear advantages and quantitative-electroencephalographic asymmetries observed during auditory stimulation; pilot testing using a combination of Neuclear Magnetic Resonance's (NMR) anatomical-imaging and chemical-spectral-analysis capabilities to study physiological activation in the human brain.

| 20. DISTRIBUTION / AVAILABILITY OF ABSTRACT | 21. ABSTRACT SECURITY CLASSIFICATION |
|---|---|
| ☒ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS | UNCLASSIFIED |

| 22a. NAME OF RESPONSIBLE INDIVIDUAL | 22b. TELEPHONE (Include Area Code) | 22c. OFFICE SYMBOL |
|---|---|---|
| Dr. John F. Tangney | (202) 767-5021 | NL |

89   3   06   117

# AFOSR 85-0379 FINAL REPORT

Judith L. Lauter, Ph.D.

Principal Investigator

## REPORT OF RESULTS

## APPENDIX

## ABSTRACT

Research interests expressed in previous publications (Lauter 1982, 1983, 1984; Lauter & Hirsh 1985) and supported under earlier grants (AFOSR 84-0335) were continued and expanded during this three-year grant period. Infrastructure requirements for the proposed research necessitated establishing laboratory facilities at the University of Arizona unique to this campus. While original plans called for a single psychoacoustic laboratory, during this project period we have organized and equipped five laboratories: 1) two psychoacoustic laboratories, 2) one sound-recording laboratory, 3) one facility for electrophysiological research in humans [both evoked potentials and quantitative EEG--shared with the Department of Neurology], and 4) a mini-computer-based laboratory currently dedicated to image processing and analysis of PET data files. [See details below under "Facilities established."]

Research activities during this period include: 1) design and implementation of microcomputer-based software to allow signal recording, generation, analysis, and automated experimental control for psychoacoustic testing; 2) data collection related to the interface between complex-sound production and perception, specifically, studies on speech acoustics, including two experiments on voice-onset-time variability in productions by speakers of several languages, and a series on acoustical characteristics of emotional expression; 3) data collection regarding individual differences in the effect of stimulus characteristics on relative ear advantages [#s 1-3 are described in detail below under "Progress on proposed and spinoff activities"]; 4) continuing data analysis and new collections documenting individual differences in auditory evoked potentials, with details related to auditory-system asymmetries; 5) preliminary tests regarding the match between behavioral measures of relative ear advantages and quantitative-electroencephalographic (qEEG) asymmetries observed during auditory stimulation; 6) pilot testing using a combination of Nuclear Magnetic Resonance's (NMR) anatomical-imaging and chemical-spectral-analysis capabilities to study physiological activation in the human brain; and 7) data analysis of Positron Emission Tomography (PET) files collected at St. Louis during 1981-1985, for evidence of individual differences in brain activation asymmetries under rest conditions, and during monaural and binaural stimulation with a variety of complex sounds [#s 4-7 are described in detail below under "Additional experiments"].

## FACILITIES ESTABLISHED

Summary. Office and research facilities for our project are distributed in four different locations around the University of Arizona campus. In the Speech Building, one of the original UA buildings, and recently awarded Historical status, are the main Speech & Hearing department offices, with secretarial support and mail service. Also in this building

is one of our two Psychoacoustic Laboratories, consisting of an 8' x 8' single-walled sound booth, computer desk and instrument rack, all housed within a large ground-floor laboratory room that contains other sound rooms used for department clinical testing and research.

Space for the PI's office and working areas for assistants are located in a small house just north of campus, within the University Expansion Area, where the University assigns housing on a temporary basis, pending purchases that will eventually lead to the clearing of existing structures to make room for new University construction. This office-house is technically off-campus, but within 5 minutes by bicycle of the Speech Building.

The Psychology Building houses three additional laboratories: the recording laboratory with an anechoic chamber, on the fourth floor, and a second psychoacoustic laboratory, and the PET minicomputer lab, housed together in a large laboratory room in the basement. Finally, the computer and subject room for electrophysiological testing are located at the University Medical Center, in the Department of Neurology. The PI assisted the Neurology department in selection of this equipment; while both equipment and space for this facility are overseen by Neurology, access for our projects is guaranteed by a new appointment for the PI to the Neurogenic Institute for Communication Disorders, directed jointly by the heads of the Speech & Hearing and Neurology departments.

Office facilities. Office support such as secretarial help, copier, package receiving, and mailboxes are located in the Speech & Hearing main offices. Work areas for PI and assistants are provided in the small house north of campus, available to the project since fall of 1986, offering approximately 800 sq. ft. of usable space. The Speech & Hearing department has helped furnish this house, with desks, chairs, file cabinets, and a typewriter, and throughout the grant period has paid for two phone lines to the house, one a UA line with WATS service, and the second a residential line, which we use for modem-based telecommunications.

Two microcomputers owned by the project are housed here for office work, a Macintosh SE with an Imagewriter II printer, and an AT&T 6300 with an IBM Proprinter.

Space in this house will support moderate expansion in terms of personnel and activities for the CNS Project (AFOSR 88-0352).

Psychoacoustic Laboratories (Speech Building; Psychology Building). The Speech-Building laboratory consists of an 8' x 8' single-walled IAC sound booth, with a desk outside for a computer and files storage, and an instrument rack. The laboratory system, state-of-the-art when purchased in 1985, is based on an AT&T 6300 microcomputer, with 20MB internal hard disk, 640K RAM, an 8087 math coprocessor, and a Data Translation 2801-A A/D board. The 12-bit board provides user-selectable gain, with a maximum of +/- 10 volts, and user-selectable sampling rate, with a maximum of 22 kHz. The interface to the board is a Data Translation screw terminal panel housed together with three anti-aliasing filters inside a box designed and built at Central Institute for the Deaf. The filters are set at

10 kHz lowpass, and serve one input channel and two output channels. When both output channels are used simultaneously, as in dichotic presentation, the available effective per-channel bandwidth is approximately 8 kHz.

Peripherals to the system include a dot matrix printer (Hewlett-Packard Thinkjet), two Crown amplifiers and three Hewlett-Packard 350D attenuators (provided by the Speech & Hearing department), and a pulse-code modulator (Nakamichi DMP-100) and video cassette recorder (Fisher 405A) for transferring live recordings (prepared on a sister system in the Psychology-Building recording lab) into the computer for processing and testing. The control computer is located on a desk outside the sound room; inside the sound room is a single-subject listening station consisting of a Zenith Z-29 terminal and AKG 141 stereo earphones.

The Psychoacoustic Laboratory in the Psychology Building has two single-walled sound booths (6' x 6' and 8' x 8'), and an AT&T 6300-based system similar to the one described above. Currently in this laboratory, subjects are tested in a "local" mode on the controlling computer's keyboard and monitor; plans are to purchase a smart terminal analogous to the Zenith Z-29 for this system, in order to separate the subject and experimenter stations.

One of the goals of the grant period was development of a software package (now called SONOS). This software serves both laboratory systems, and represents a range of functions from signal generation to automated testing. Component programs are: SERIES (generation of pure and complex tones using either Fourier or cosine specification formats, with adjustable amplitude, frequency, and phase for a maximum of 72 components); RECORD (for recording external sounds into the system); EDIT (capabiliities for a variety of waveform display, replay, editing, and sequencing functions); LIST (ability to gather sounds edited and labelled in EDIT into lists which can be called by the experimental program, with each list identified by name, key-coding per sound, and masker for dichotic presentation); and TRIAL (automated experimental control, with user-selectable sound sets, testing intervals, presentation mode [right, left, dichotic], ear scored, and number of trials per block; the program formats each block, times all intervals, scores responses, and presents feedback after each trial and after each block).

Capabilities provided by these laboratories have made possible our research projects on speech acoustics (sound analysis) and dichotic listening (sound synthesis, recording, experimental control), to be described later in this report. In the future, the laboratories will provide facilities for a variety of psychoacoustical tests, including those required for the CNS Project.

Recording Laboratory. This laboratory makes use of an anechoic chamber (12' x 12') which was designed into the original 1965 blueprints for the UA Psychology Building, but which had not been used except for storage before our project arrived on campus in 1985. Inside the chamber the subject station includes a comfortable chair, an AKG 451E microphone on a mike stand, and a General Radio 1562-Z sound-level meter (SLM

provided by the Speech & Hearing department).  Outside the chamber, the experimenter
station consists of a chair and desk, with a PCM/VCR system (Sony PCM and Mitsubishi
4-head VHS VCR) providing low-cost high-resolution digital recording.  Signal analysis of
sounds recorded in this laboratory is accomplished using our AT&T 6300 machines
running our SONOS software, and/or a Kay 7800 digital sonograph, owned by the UA
Linguistics Department's Phonetics Laboratory.

The Recording Laboratory provided facilities used in our research projects on speech
acoustics, including speech-cue variability and emotional expression, and for collecting
real-speech tokens for dichotic testing.

PET Data-Analysis Laboratory.  Equipment for this laboratory was purchased under
AFOSR 87-0003, and includes:  a Perkin-Elmer 3205 minicomputer, a Ramtek MC68000
Color Display Controller and 19″ RGB monitor, and a Matrix 3000 film recorder.
Although this is a general-purpose laboratory, it is currently dedicated to analyzing data
collected by the PI during 1981-1985 at the Washington University Mallinckrodt Institute
of Radiology's PET-scan laboratory in St. Louis.  Our data library consists of complete
files representing  a 31-study series involving auditory stimulation in 17 normal young
adults, and is being analyzed using programs included in a software package developed at
St. Louis and provided free to our UA laboratory.

Work is currently in progress on this system directed to examining the degree of
hemispheric asymmetry observed in PET scans of brains both at rest and under stimulation
conditions with a variety of presentation and stimulus designs (see below).

Human Electrophysiology Laboratory.  The PI acted as a consultant to the Department
of Neurology during the selection process for equipping this laboratory.  The system
purchased is a Brain Imager made by Neuroscience, and offers a variety of testing options,
including evoked-potential testing sampling several levels of the CNS via several sensory
modalities, as well as multi-channel brain mapping based on ongoing EEG.  This machine
has just been installed;  for the experiments on qEEG and auditory brainstem responses
(ABRs) described below, we arranged for use of a Cadwell Spectrum 32 in a private-
practice office in Tucson (for qEEG), and a Nicolet 2000 owned by the University Medical
Center's Department of Otolaryngology (for ABRs).

## PROGRESS ON PROPOSED AND SPINOFF ACTIVITIES

Summary.  Included in the original proposal for this grant were plans to:  1) Establish a
laboratory for dichotic testing, including development of software with a variety of
capabilities;  2) Collect real-speech samples of stop-consonant-vowel (CV) productions
from a number of languages, analyze them to determine timing characteristics, and prepare
tokens for dichotic testing;  3) Test individuals from each of the language backgrounds
sampled in #2 to determine relative ear advantages for the stop CV tokens;  4) Prepare a

series of complex-sound sets for use in a group of experiments on the effect of stimulus characteristics on ear advantages evoked from English-speaking listeners; and 5) Test a group of subjects on these sound sets.

Goal #1 (establishment of a laboratory) has been completed; hardware details are given above under "Facilities." The software developed for this project, comprising our package called SONOS, is also described above.

Goal #2 (collection, analysis, and processing of real speech CVs) has been completed. Data collection made use of the Recording Laboratory, and analysis and processing were accomplished via our SONOS software. In addition, the analysis stage of this project generated a spinoff study of VOT variability, which has resulted in three meeting presentations, one MS in review, and one MS in preparation. Also, interest in segmental speech timing led to a complementary project regarding the acoustical characteristics of suprasegmental aspects of speech, specifically, emotional expression. Descriptions including sample data from all of these original and spinoff studies are given below.

Goal #3 (dichotic testing of multi-language stop CVs) could not be pursued. After tokens were edited from the original recordings and prepared for testing (both functions accomplished using our SONOS programs), pilot listening (also via SONOS) by four subjects indicated that these real-speech tokens were "too easy" to be useful for dichotic testing: i.e., even with dichotic presentation, subjects could quickly achieve ceiling performance attending to either ear. Multi-language testing of listeners from other-than-English language backgrounds will have to wait on appropriate stimulus design.

Goal #4 (preparation of complex-sound sets for dichotic testing with English-speaking listeners) has been completed, using our SONOS software. Included in our sound library currently are sets of pure-tone three-note melodies, with a variety of inter-onset-timings and pitch steps, sets of melodies made with complex tones where melody is determined by spectral changes in highlighted harmonics, and sets of melodies made with noise bands. Descriptions are given below.

Goal #5 (testing with sounds from #4) is in progress. Due to delays in programming, dichotic testing on the sets of complex sounds was not started before late summer of 1988. Results of testing to date are given below.

"Speech dictionary" collections. It was decided that as a context for the proposed collection of stop-CV productions from four languages (American English, Japanese, Mexican Spanish, Navajo), complete sets of the phonemic "syllabic building blocks" from each language would be collected from each subject; and, in addition, that we would collect these sets in a repeated format, to build a database documenting intra-individual as well as inter-individual variability in phoneme production.

The building block elements to be collected were: 1) syllable-initial consonants, 2) syllable-final consonants, and 3) syllabic nuclei. For the first three languages named above, members of each set of elements were included in an appropriate context (e.g, "ha-Ca" for syllable-initial consonants, "haC" for syllable-finals, and "hVd" for syllabic

nuclei), to be repeated six times by each talker. For Navajo, we were restricted to the use of real Navajo words due to subjects' lack of familiarity with written Navajo. As a result of vocabulary limitations regarding consistent phoneme combinations (e.g., each of the syllable-initial consonants followed by a low back vowel), we were unable to collect a complete phoneme library for Navajo, and prepared only lists of real Navajo words representing each of five Navajo stops in initial position followed by a low back vowel.

Separate lists of syllable-initial stop-consonant tokens were prepared for each of the languages except Navajo, so that acoustical characteristics of the stops produced in the context of only stops could be compared with characteristics of stops produced in the context of all other phonemes of the language. These isolated stop-consonant sets were also to be read six times by each talker.

Three female speakers of each of the four languages served as subjects. As data collection proceeded, another graduate student's interest in the project resulted in the addition of Mandarin Chinese to the library. The Mandarin Chinese collections were more extensive than those from the other languages: 1) productions by three female and three male talkers were recorded; 2) for the isolated stop-CV tokens, each Chinese talker began by pronouncing the six stops as English sounds, reading them from a list cued with English words, and then went on to do the Mandarin stop series; 3) for the Mandarin versions of the stops, each talker produced one complete repeated series (six repetitions) of the set of stops produced with each of the five Mandarin tones.

After recordings were complete, one set of productions from one talker from each of the five languages was used to produce a "speech dictionary" of spectrograms illustrating the phonemes of that language, with sections for syllable-initial and syllable-final consonants, and a section for syllabic nuclei. Talkers were offered copies of these dictionaries as compensation for their participation in the project. Sample spectrograms for six English stop-CV syllables and five Navajo stop-initial words are reproduced in the Lauter & Pearl (1986) presentation text included in the Appendix.

VOT variability. The productions of "isolated stops," where the lists of syllable-initial stops were repeated six times served as a database for analysis of intra-individual variability in VOT timing. Using the SONOS software, each production of a stop was recorded into the computer, its waveform displayed, and cursors set to select the VOT portion of the syllable, which was then printed out onto a dot matrix printer. These printouts, showing the VOT portion of the six repetitions of each stop produced by each talker were combined onto a single page, and used to measure VOTs.

Figure 1 presents a set of six such waveforms for one talker's productions of "ga," with VOT offset marked with an arrow. All talkers showed a variation in the relative durations of VOTs for the six stops, from set to set. Figure 2 presents absolute VOT durations for all tokens for all six repetitions by another talker. Note the fluctuation in relative duration, such that no repetition comprises the regular series of VOT increasing

from bilabial to velar, voiced to voiceless, that would be predicted based on the synthetic-speech literature.

Calculation of mean VOTs for all talkers revealed a language-specific characteristic: American English, Mandarin Chinese, and Navajo make a distinction between two groups of stops based on VOT, while Japanese and Mexican Spanish do not (examples of mean VOTs for one talker from each language shown in Fig. 3, Panel A). Our mean VOT values for English and Mexican Spanish agree with those previously reported; our measurements for Japanese, Mandarin, and Navajo are original.

Standard deviations varied from talker to talker, with no apparent pattern or language specificity (see Fig. 3, Panel B for VOT s.d.'s for the same five talkers of Panel A). Finally, a ratio of these two measures (s.d./mean), the Coefficient of Variation (V), revealed not only individual differences (Fig. 3, Panel C), but for all talkers studied, indicated that there may only be a restricted number of such patterns, which seemed to be language-indepenent (Fig. 4 shows the patterns observed thus far).

The relative variability patterns (i.e., values of V: Fig. 4) may prove to be informative regarding production constraints and/or "styles," which are context-sensitive. This conclusion was suggested by the extended Mandarin Chinese study, where we were able to observe VOT variability as a function of a change in context: namely, the fundamental frequency pattern of the pronounced token. Although all of the tokens intended as Mandarin stops were produced with different F0 patterns, it happened that the female Mandarin speakers produced their tokens intended as English stops with the same F0 contour as their Mandarin tone-4 tokens. This made it possible for us to compare the variability patterns for two sets of stops produced in the same context (tone-4-type falling F0 contour, used for Mandarin tone-4 and the English tokens) and variability patterns for stops produced in different contexts (Mandarin tones 1, 2, 3, 4/English, and 5).

Figure 5 presents schematics of the F0 contours used by one Mandarin Chinese female talker for tokens produced with the five different Mandarin tones, and English. Note the similarity between her tone-4 and English contours. All three female Mandarin speakers showed very good matches between tone-4 and English VOT V patterns (Fig. 6, Panel A), and frequency-contour characteristics (Fig. 6, Panel B). Analysis of the male talkers' F0 contours is currently in progress, to determine whether there are also instances of context-sensitivity in their VOT variability patterns.

These results may have implications for studies of speech production, in that they suggest a potentially useful new means of studying interactions between different articulators, and the ways in which such interactions bear on issues related to speech-planning stages and mechanisms. For example, for any one talker, are there patterns of complementary speech-cue variability? For example, does a talker who is very inconsistent about VOT timing for /t/ show a high consistency in another cue for /t/, such as burst amplitude? Are there predictable trade-offs in cue variability, such that consistency in one cue "allows" inconsistency in one or more other cues? Are the rules governing which cues

may interact in this way related to the phonology of different languages? For example, for our Mandarin talkers, VOT and F0 in these tokens were both segmental cues—would English speakers show the same degree of match between VOT variability pattern and F0 contour, even though VOT for them is a segmental aspect, and F0 contour a suprasegmental aspect, of speech?

These results regarding speech acoustics are also clearly important for considering how the auditory system functions in speech perception. Our specific interest is in the temporal aspects of this particular situation, where a short-time-base characteristic (VOT, with distinctions often based on a few msec) seems to be interacting with a long-time-base characteristic (F0 contour, with distinctions based on patterns that are articulated over many tens of msec). Eventually our serendipitous observation of a possible VOT/F0 interaction may lead to hypotheses regarding brain mechanisms for speech perception, and their relation to mechanisms for the perception of complex sounds in general. Although differing cerebral organizations for "phonetic perception" vs. "emotional perception" have long been posited, the idea that such distinctions could be based on physical characteristics of test stimuli has yet to be generally accepted. In future stages of our CNS Project, it might be possible to view brains during stimulation with sounds which tap the different posited mechanisms: real speech and music with different acoustical characteristics and emotional associations, synthetic complex sounds designed to mimic these various aspects of actual sounds. Of course, there are also implications for asymmetries of cerebral organization, with regard to variables of emotion vs. non-emotion, positive vs. negative emotion, and rapid vs. slow sequential timing.

This research has been reported in three meeting presentations (Lauter & Pearl 1986; Lauter & Lu 1987 a,b), and is described in one MS currently in review (Lauter et al, submitted to JASA) and one MS in preparation (Lauter & Lu). (See Appendix for some texts.)

Acoustics of emotional expression. The importance of emotional expression as an interface between acoustical characteristics of speech on the one hand and brain perceptual organization on the other served as the motivation for our second spinoff project on real-speech acoustic phonetics comprising a Ph.D. dissertation by C. Baldwin co-directed by the PI and M. Wetzel of the UA Dept. of Psychology.

Although there is a growing literature regarding brain asymmetries for the identification of emotions, there are few systematic studies on the acoustics of emotional expression in speech. For our study, six professional actors, and six speakers with no training (three females and three males in each group) produced each of two sentences, "Of course I love you" and "The horse tries one food," with a series of emotional colorations: neutral, angry, sad, afraid, surprised, happy, disgusted. Digital recordings of all productions were made in our Recording Laboratory, and tokens were analyzed using the Kay 7800 digital sonograph, and our SONOS software.

Sample waveforms with superimposed amplitude contours of the sentence "Of course I love you" spoken in a neutral, happy, and sad tone of voice by a male actor are presented in Fig. 7. Measurements to date on the library of productions include: 1) amplitude profiles, describing the amplitude contour of the sentences in terms of five points, at peak amplitude of each of the five syllabic nuclei (cf Fig. 8, Panel A); 2) amplitude mean, range, and standard deviation characterizing each emotion produced by each group of talkers; 3) rhythm profiles, describing the sequence of syllable durations for the five syllables; and 4) syllable-duration mean, range, and standard deviation for each emotion produced by each group of talkers (cf. Fig. 8, Panel B).

Continuing analyses of the same library include: 1) complete fundamental-frequency (F0) contours, with descriptions in terms of maximum and minimum and rate of change; 2) fundamental-frequency profiles, based on the F0 value at peak amplitude of each of the five syllabic nuclei; 3) F0 mean, range, and standard deviation for each emotion produced by each group; and 4) segmental and sub-segmental analysis, examining how characteristics of amplitude, duration, and spectrum for each acoustical event in the sentence change from emotion to emotion. A more complete picture of how emotional expression is cued in speech will help us design sounds for psychoacoustic and physiological testing, to explore the importance of different types of complex-sound cues for perception and for determining the organization of CNS response.

This work has been reported at two meetings (Baldwin et al 1988a,b), and results to date are described in two MSS in preparation (Baldwin & Lauter, and Lauter & Baldwin).

Dichotic listening. Sound libraries created with our resources including the Recording Laboratory and the SONOS software consists of:

1. American English stops (ba,da,ga,pa,ta,ka spoken by
      female talker; approx. 250 ms/token)
2. Japanese stops (similar set)
3. Mexican Spanish stops (")
4. Mandarin Chinese stops (")
5. Navajo stops (ba, da, ga, t'a, k'a spoken by female talker)
6. synthetic versions (Haskins-Lab-type) of English stops
      with /a/
7. WT200, 100, 50: three-pure-tone melodies made with
      whole-tone steps, centered at 1480 Hz; the three
      sets are timed at 200, 100, and 50 ms inter-onset-
      interval (IOI)
8. HT 200, 100, 50: same as #7, but with half-tone steps
9. H/QT 200, 100, 50: same as #7, but with a frequency
      step halfway between quarter and half

10. QT 200, 100, 50: same as #7, but with quarter-tone
    steps
11. ET 200: same as #7, but with "eighth-tone" steps (20
    Hz at this frequency)
12. BB noises 200, 50: three-noise melodies made with
    1000-Hz-wide ("BroadBand") noise bands created
    using SoundEdit running on a Macintosh SE; bandpass
    values for the low noise = 500-1500 Hz, for the mid
    noise = 1000-2000 Hz, for the high noise 1500-2500
    Hz; noise onsets are timed at 200 ms IOI for one set
    and 50 ms IOI for the second
13. NB noises 200, 50: three-noise melodies made with
    500-Hz-wide ("NarrowBand") noise bands created
    using SoundEdit running on a Macintosh SE; bandpass
    values for the low noise = 1000-1500 Hz; for the mid
    noise = 1500-2000 Hz; for the high noise = 2000-2500
    Hz; noise onsets are timed at 200 ms IOI for one set
    and 50 ms IOI for the second.

All of these sounds have been created and stored on disk both as single sounds, and grouped into sets with the indicated names. For testing, sound sets can be called via their group name (e.g., "English stops," "WT200"), a call which also provides the TRIAL program with the response keys encoded for each sound in the set.

In preparation are melodies made with complex tones, to be created with the SERIES program of our SONOS software.

As noted above, the real-speech stop-CV sets were shown in pilot testing to be too easy for use in dichotic listening tests: even with dichotic presentation, these real-talker sounds are so rich with redundant cues to stop identity that listeners can achieve ceiling performance attending to either ear. Thus the originally proposed experiments testing listeners from various language backgrounds on their own and each other's stops will have to be postponed until we can design more useful stimulus sets.

Testing has begun on the other test sounds. Data for the hypothetical "extremes," the synthetic stops and the 200-ms pure-tone patterns are now available for a total of five English-speaking listeners, and for two native speakers of other languages: one Japanese male and one Mexican-Spanish female. Results for the seven subjects are displayed in Fig. 9, in the form of a Relative Ear Advantage (RelEA) plot (cf. Lauter 1982, 1983, 1984), with one row for each listener, and ear-advantage scores indicated for each listener tested on each sound, EAs for "200T" indicated by filled circles, and EAs for the synthetic CVs ("sCV") by open squares. Some of the listeners have been tested on a third set, of 50-ms tone patterns; EAs are indicated by the open circles. Procedures used in this testing are the

same as those described in our earlier publications, except accomplished on our AT&T 6300 microcomputers running the SONOS software; the EA values plotted are based on 216 trials per ear of report.

Sound-set series to be used in continuing testing will enable us to observe changes in EAs as a function of: 1) timing alone (200, 100, and 50-ms pure-tone patterns); 2) pattern bandwidth alone (100-ms-timed tone patterns differing from set to set only by the pitch-step used in the pattern, whether within one whole tone, between whole tone and octave, and larger than octave); 3) element bandwidth alone (100-ms-timed patterns made with pure tones, complex tones, or noises); 4) interactions between these dimensions, combined in pairs (two combinations to be tested) or three at a time.

Extension of these data in the CNS Project will involve testing Project subjects on three of the library's sound sets which should evoke a triad of distinctive EAs: synthetic stop-CV syllables (for a hypothesized extreme rightward EA), 200-ms-timed pure-tone patterns (for a hypothesized extreme leftward EA), and 50-ms-timed pure-tone patterns (for a hypothesized mid-position EA). Subsequent to determination of the EAs using behavioral testing, each subject will then be examined using a number of physiological brain-scanning methods (Positron Emission Tomography PET, Magnetoencephalography MEG, quantitative Electroencephalography qEEG, etc.) while being stimulated with the same sets of test sounds, to enable us to observe the relation between physiological brain asymmetries derived for each subject and the EAs measured using behavioral methods.


## ADDITIONAL EXPERIMENTS

Summary. In parallel with the proposed and spinoff activities, we have been able to make use of previously-collected data, as well as to establish collaborative arrangements for access to equipment for collecting new data, regarding aspects of brain physiology that may underlie the phenomena observed in our behavioral tests of asymmetries in complex-sound perception. These experiments represent extensions of earlier research by the PI (e.g., brain electrophysiology: Lauter & Loomis 1985; PET: Lauter et al 1983 a,b, 1984, 1985 a,b).

Evoked potentials. Our observations of individual differences in auditory asymmetries as expressed in dichotic tests led us to look for physiological correlates. Initial interest focused on evoked potentials (EPs), tested under a repeated-measures experimental design. Although a number of researchers had studied between-subject variability (see Lauter & Loomis 1986a for a review), there was very little information regarding within-subject variability for EPs at any level of the CNS.

Our activities related to EPs during the grant period included: 1) continuing analysis of a repeated-measures auditory EPs data collection on 7 adults, conducted during 1983/84 under a collaborative arrangement with the Washington University Department of

Otolaryngology; and 2) new data collection and initial analysis of test series also using collaborative arrangements, for three new groups of subjects: eight more young adults, 5-year-olds, and 10-year olds. Findings from all of these studies are described below.

Analysis of old data: During 1983/84, in a collaborative arrangement with the Washington University Dept. of Otolaryngology in St. Louis, data collection was completed on a group of 7 audiologically and neurologically normal young adults, tested for both brainstem and cortical auditory EPs under two monaural as well as binaural stimulation conditions. Each subject was tested in a series of eight weekly sessions, on the same day of the week and same time of day for each subject. Analysis of the latency variability in these data, for both brainstem and cortex, was reported in a meeting presentation (Lauter & Loomis 1985). During the project period, the brainstem latency findings were published (Lauter & Loomis 1986a: see Appendix).

These analyses of repeated-measures EP data suggest that the variability of EP parameters such as peak latency may provide a much more sensitive index of subject characteristics as well as of determinants of sensory-system response such as ear-of-presentation than do absolute values of EP parameters, the conventional means of describing EPs. The difference in the amount of information contained in absolute EP latency values vs. measures of EP variability is illustrated in Fig. 10, Panel A. On the left are plotted mean latency values for each of five ABR peaks, averaged across seven listeners and eight sessions per subject. Note that, as is well known, the mean-latency functions for all three ear conditions are virtually identical. On the right are plotted values for an index of the variability of these peak latencies (the index is the reciprocal of Pearson's Coefficient of Variation—in this form, mean/s.d.), calculated separately for within- and between-subject comparisons of responses under the three stimulation conditions: right-ear, left-ear, and binaural. Note that, unlike the mean absolute values on the left, the variability data in the right panel show distinctions based on within- vs. between-subject comparisons, as well as differences due to ear-of-presentation.

Similar "stability profiles" can also be calculated for individual subjects, comparing responses to different ears, or response patterns during the two halves of the experiment (first 4 weeks vs. second 4 weeks). Figure 10 Panel B presents such profile comparisons for four subjects. Note that for these ears tested in these subjects, there is good replicability of the stability profile.

The next phase of analysis of the data collected in the 1984 series, analysis conducted during the grant period, focused on amplitude of the five ABR peaks. Although it is well known that EP peak amplitudes are much more variable than EP peak latencies, our analysis of the repeated-measures data indicated that even so, calculation of ABR peak-amplitude variability reveals unsuspected patterns. For example, Fig. 11, Panel A presents a parallel to Fig. 10, with absolute mean amplitudes plotted on the left, contrasted with amplitude variability displayed on the right. The absolute-value panel shows the well-known difference between EP amplitudes for binaural vs. monaural stimulation; the

variability panel also shows this difference, along with a distinction based on within- vs. between-subject comparisons.

Amplitude stability profiles can as well be calculated for individual subjects. Figure 11 Panel B presents some examples, including an instance of a characteristic observed in both the latency and amplitude data--at times, a given subject's stablity profile may require the second 4 weeks of testing to acquire the shape which another subject's profile shows consistently in both months (cf. KP's <Binaural-Right> profile which takes two months to resemble SA's consistent <Binaural-Right> profile).

Amplitude stability profiles may also reveal asymmetries in ABRs, perhaps related to findings obtained with other methods, such as the version of the Binaural Interaction Component (BIC) described by Berlin and colleagues (e.g. Berlin et al 1984), based on waveform addition and subtraction. For example, Fig. 12 compares an ABR amplitude-stability profile set indicating a clear asymmetry favoring right-ear input (top panel, from our ABR data) with a derived waveform illustrating a right-ear BIC (lower panel; the BIC is the interpetation of the positive-going peak at 5 ms--adapted from Berlin et al 1984).

The amplitude-variability data have been reported in a meeting presentation (Lauter & Loomis 1986b) and in a publication (Lauter & Loomis 1988).

New data series #1: During 1985/86, a collaborative arrangement with the Pediatric Audiology section of St. Louis Children's Hospital resulted in data collection on an additional group of 8 audiologically and neurologically normal young adults. These subjects were tested for ABRs and middle-latency responses (MLRs) to monaural and binaural stimulation, collected in eight weekly sessions per subject. Besides providing a replication of our earlier observations on ABR variability, these data allowed us to compare patterns of ABR and MLR variability. Figure 13 Panel A presents a graph parallel to that shown in Fig. 10 Panel A, comparing absolute mean latencies for five MLR peaks (three vertex-negative and two vertex-positive, on the left) with MLR latency variability for the same five peaks (on the right). Note that the variability functions show distinctions between comparisons (within- vs. between-subjects) as well as between ears, that are not visible in the absolute data. As with individual subjects' ABR latency-variability profiles (cf. Fig. 10 Panel B), individual MLR latency-variability profiles can also show good replication from month one to month two (Fig. 13 Panel B).

This new database makes it possible to compare ABR and MLR variability. Figure 14 Panel A presents a summary of the variability calculations for this new group of eight young adults, with ABR curves on the left and MLRs on the right. The striking difference between the within-subject stability of MLR peak No and the later MLR peaks led us to a preliminary comparison of variability data for ABR, MLR, and cortical responses. Figure 14 Panel B illustrates this comparison, based on ABR and cortical data for our first group of adults, and MLR data from the second group. Note that the variability of MLR peaks Po through Nb are in the same range as the cortical values, while MLR peak No variability is

intermediate between that of ABR (more consistent) and later MLR peaks (more variable). This distinction between MLR peak No and other MLR peaks has been noted by other researchers using other methods, and it has been speculated that the observed characteristics indicate that the generators of MLR peak No are located in the brainstem, while nuclei important for later MLR peaks are located in the thalamus and above.

The latency findings for this second group of adults, with comparisons between ABR, MLR, and cortex, were reported in a meeting presentation (Lauter & Karzon 1987), and two MSS based on data from all 15 subjects are in preparation (Lauter & Karzon, In prep. a,b).

New data series #2: A PhD dissertation project by J. Lord-Maes, co-directed by the PI and S. Mishra of the Educational Psychology Department at UA, represented the first extension of our repeated-measures EP testing, to another age group. Seven children between 5 and 7 years old, four females and three males, were tested in a design similar to that used with the young adults: eight weekly sessions per subject, with monaural and binaural stimulation. Only ABRs were collected, and monaural responses were stored using both ipsilateral and contralateral references, to enable us to study the variability of derived-waveform BICs, as well.

Figure 15 Panel A presents our basic calculations for this new group of subjects, contrasting absolute ABR latencies (on the left) and ABR latency variabilities (on the right). Note that the variability profiles reveal the same types of distinctions between comparisons, and ears, as observed in the 15 adults. Although in these younger subjects, replicability of individual profiles was rarer than in the adults, some ears in some subjects did show good consistency from month to month (Fig. 15 Panel B).

Analysis of these data is continuing, with plans to study in more detail the patterns of individual differences, amplitude variability, and variability in BICs generated by adding and subtracting waveforms according to the method described by Berlin et al (1984). We also intend to compare values and patterns observed in this group of subjects with those seen in our adults, as well as from the subjects from the next experiment, described below. A summary of these findings will be presented to the Acoustical Society meeting in spring 1989 (Lord-Maes & Lauter 1989).

New data series #3: A PhD dissertation project by R. Oyler, co-directed by the PI and N. Matkin of the UA Speech & Hearing Sciences Department, focused on repeated-measures ABRs in an older group of children. The nine subjects were between 10 and 11 years old, and were all males. This experiment involved a slight change in the design used previously, reducing the overall time of the test series from two months to one, but collecting six waveforms per ear of presentation from each subject during each of the four weekly sessions.

Comparison of the mean absolute latencies and the latency variability is presented in Fig. 16 Panel A (calculations based on values for peaks in the first ABR waveform collected in each of the 4 weekly sessions, to be most comparable with Figs. 10 and 15).

Data documenting profile replicability within subjects comparing month one and month two are of course not available in this experiment. However, profiles calculated for the six waveforms of the first session and the six waveforms from the final session show good replicability for some ears in some subjects (Fig. 16 Panel B).

With these three sets of experiments, with each focusing on a different age group, we are able to compare the degree of within- and between-subject consistency seen for 5-6-year-olds, 10-11-year-olds, and young adults. Figure 16 Panel A makes such a comparison; these graphs suggest that although there is little change in between-subject consistency from age group to age group, there are clear changes in the within-subject consistency--note that the 5-6-year-olds are almost as unlike themselves (WS curves) as they are unlike each other (BS curves). These distinctions provide yet another indication of the greater sensitivity of ABR variability measures as compared with absolute parameter values: ABR absolute peak latencies for all three groups are the same (clinically, it is said that a 3-year-old's ABR is the same as an adult's)--but the stability curves clearly suggest that ABR variability may be sensitive to developmental changes that go on long after the ABR absolute peak latencies reach their "adult" values.

Another interesting comparison provided by the design of the experiment with the 10-year-olds is between the degree of variability seen in repeated ABRs collected within the same hourly session, and ABRs collected on a weekly basis. Figure 17 Panel B presents such a comparison, showing within-subject profiles calculated either: (on the left) for the first waveform collected on each of the four weekly sessions (most parallel to the design used for the other two age groups), or (on the right) for all five waveforms collected on each of the four test days. Note that there is a clear distinction in the degree of variability based on whether waveforms were collected sequentially in the same session, or are taken at weekly intervals: within-session measures yield much higher within-subject consistency values.

These differences have implications for future use of repeated-measures EPs. The comparisons shown in Fig. 17 Panel B suggest that for group studies, multiple ABRs collected during the same session may be very useful for illustrating differences between ears--cf. the clearer ear differences shown here in the 5 runs/day vs. the 1 run/day graph. However, individual profiles as represented in Fig. 16 Panel B suggest that for individuals, multiple ABRs collected during the same session may be "too consistent," i.e., our variability index reaches ceiling levels, and thus may not be useful for studying characteristics of EPs such as individual characteristics, maturational changes, or asymmetries, on a subject-by-subject basis. More extensive tests, comparing ABR variability based on different collection schedules in the same subjects, are planned for the future.

Further analysis of the 10-year-old data will include more detailed examination of individual differences, especially comparing the two test schedules, amplitude variability,

and occurrence and variability of asymmetries, measured using both amplitude-variability comparisons and derived-waveform BICs.

Presentation of these data is planned for the fall 1989 Acoustical Society meeting (Oyler & Lauter 1989).

All of the work on repeated-measures EPs will provide a background for our use of EPs in the CNS Project, where subjects will be tested to determine ABR asymmetries (in responses to the same types of stimulation used in the experiments described above, 11/sec filtered clicks), and asymmetries will be quantified in terms of direction and magnitude, based on amplitude variability as well as BIC calculations.

qEEG. Interest in identifying physiological correlates of behavioral relative ear advantages led us to explore the possibility of monitoring ongoing EEG collected from subjects while they were being stimulated with the same types of sound presentation used in our dichotic-listening experiments: monaural vs. dichotic with directed attention, complex-sound A vs. complex-sound B, etc. A collaborative arrangement with the Brain Mapping Laboratory in Tucson allowed us to test a pilot series of four subjects.

Subjects were pre-tested using behavioral methods, to familiarize them with monaural and dichotic identification of two sound sets: 1) six synthetic stop consonant-vowel syllables (coded as "S" in the figures below), and 2) six three-note patterns made with three pure tones set at 1440, 1480, and 1520 Hz, with the 25-msec tones set at 200 ms between onsets (coded as "T"). Procedures standard in our laboratory for dichotic testing were used.

For two of our subjects, who were trained listeners, complete test series were collected for both sound sets (Subjects JL and CB). The other two subjects (DW and JM) found dichotic identification of the tone patterns quite difficult, and were unable to remain in the experiment long enough to achieve better-than-chance performance with dichotic presentation.

After behavioral testing, individuals were scheduled for a qEEG session. Preparation included fitting of an electrode cap, with leads connected to a Cadwell Spectrum 32 qEEG system, with capabilities for multi-channel data collection and spectral analysis. Electrodes were placed at 8 locations over each hemisphere, and 5 locations along midline, according to the 10-20 system; potentials at all locations were referenced to linked earlobes. When impedance for each of the leads was less than 8 ohms, testing was begun.

The schedule of conditions is shown below. A time base similar to that used in the behavioral testing is used, with 5 min of EEG collected during each test condition. Note that each qEEG test session concludes with a brief set of blocks involving motor activation. Throughout, the subject reclines in a comfortable chair in a quiet, darkened room. Test sounds are played via a stereo cassette recorder through stereo earphones. During monaural stimulus conditions, subjects are told to attend to the ear of input; during dichotic conditions, they are told to attend to the ear targeted for that condition in the same way done for the behavioral tests previously. We do not ask for score-able identification performance

during the EEG testing, in order to avoid movement artifacts. Trained subjects report that it is easy to perform this "mock" dichotic listening. The qEEG results suggest that the two trained listeners here were in fact successfully replicating processing patterns used in the behavioral testing.

qEEG conditions tested per session:

1. Control (no activation)
2. Synthetic syllables in left ear
3. Synthetic syllables in right ear
4. Synthetic syllables dichotically, attend to right ear
5. Syllables dichotically, attend to left ear
6. Control
7. Tone patterns in right ear
8. Tone patterns in left ear
9. Tone patterns dichotically, attend to left ear
10. Tone patterns dichotically, attend to right ear
11. Control
12. Preferred hand flexion, 1/sec
13. Opposite hand flexion, 1/sec
14. Bilateral hand flexion, 1/sec
15. Control

Data were analyzed off-line. From each 5-min EEG record, 36 2.5-sec artifact-free epochs were selected by eye. The Cadwell Spectrum 32 then averaged the selected epochs, performed spectral analysis according to 4 EEG bandwidths, and displayed the results in terms of a number of parameters (power, power asymmetry, coherence, etc.). From each such set of values representing each subject tested on each condition, a single number was chosen for analysis: 1) for the auditory conditions, the beta power asymmetry comparing temporal-lobe electrode locations T3 and T4 was selected; and 2) for the hand-movement conditions, the beta power asymmetry comparing F7 vs. F8.

Figure 18 presents qEEG results for the motor activation conditions tested on subject JL, with beta power asymmetry at electrode locations F7/F8 plotted on a "relative hemisphere advantage" graph analogous to our relative ear advantage graph (cf. Fig. 9). Note that the results for this subject: 1) show a clear contralateral activation pattern for unilateral movement, with a left-hemisphere advantage for right-hand movement ("HR"), and vice versa; and 2) may provide evidence of a higher-order principle of asymmetry, perhaps related to interhemispheric coordination, as the both-hand movement evoked an symmetry almost identical with right-hand movement (the preferred hand for this subject).

Two other motor conditions were tested on this subject: tongue-tip movement (in the vertical plane), and respiration during silent speech (no laryngeal or upper-articulator movement). Note that these two speech-related motor conditions evoked very different power asymmetries, with respiratory control ("Resp") showing no asymmetry at F7/8, and tongue-tip movement ("Tongue") evoking the largest left-hemisphere dominance seen in this subject for any condition tested.

Two of the four subjects tested showed similar resemblances between bilateral hand movement and movement of one of the other hands (R for JL, L for CB). Failure of the other two Ss to show such a match may be due to the high levels of artifact present throughout their records during these conditions, which were tested late in each session. In the future, we plan to test the somewhat fatiguing hand-flexion conditions first, while the subjects are fresh.

Results for all auditory conditions as tested on subject JL are presented in Fig. 19. The T3/T4 beta-asymmetry values representing control, monaural, and dichotic conditions are plotted on separate rows. The qEEG results for JL and the other three subjects indicate evidence of two types of auditory-response asymmetry: 1) one based on "side of space" in that attention to right vs. left ear results in opposite asymmetries; and 2) an asymmetry based on type of sound, in that attention to syllables vs. tones results in opposite asymmetries. There are also interactions between the two types of asymmetry, such that right-ear syllables tend to evoke one extreme of asymmetry and left-ear tones the opposite extreme. Note, however, that none of the auditory conditions evoked an asymmetry favoring the left hemisphere in this subject--changes in asymmetry from condition to condition are articulated in terms of modulations of RHA.

This observation is in contrast to the behavioral results for subject JL, as shown in Fig. 20, where comparisons between the EAs tested behaviorally and "hemisphere advantages" (HAs) calculated for the qEEG results, for each of the four subjects, are plotted. An example of the procedure used to calculate the qEEG HAs is given below the figure. Behavioral testing of subject JL resulted in the syllables evoking a 20% right-ear advantage, and the tones a 20% left-ear advantage. The qEEG asymmetries suggest that JL's behavioral results may indeed reflect changes in relative hemisphere activation, but these seem to be changes which occur in the context of a continuing processing predominance favoring the right hemisphere.

The comparison presented in Fig. 20 between relative EAs and relative HAs for all four subjects indicates that although there may be inconsistencies in the absolute dominance for any one sound (e.g., JL and CB show right-hemisphere qEEG advantages for the syllables although their behavioral REAs would predict left-hemisphere dominance for these sounds), there is good agreement in the relative patterns of asymmetries comparing the two sound sets under the two measures. Note also that subjects JL and CB show the predicted "right-hander" pattern of syllables evoking EAs that are to-the-right of those for the tone

sequences, and HAs that are to-the-left, while the results for subjects JM and DW (the only left-hander in the group) suggest a reversal of this pattern. A more complete description of data for all subjects is presented in Lauter (1988a), included in the Appendix.

Several questions are generated by these preliminary data; however, we believe that the results are encouraging regarding the potential usefulness of qEEG as a tool for studying cerebral responses to fairly simple stimulus and task combinations, and indicate that "cognitive" processes are not the only phenomena that might be usefully studied using qEEG. The degree of individual differences seen in all of these results suggests that more subjects need to be tested, particularly if we are to understand the significance of the reversed patterns of activation shown by some of the subjects. Future qEEG experimental designs will also require all subjects to complete behavioral testing on all sounds before physiological testing.

These results will assist us in designing the qEEG test series to be included in the CNS Project. Testing for this project will be conducted using the Dept. of Neurology's Neuroscience Brain Mapper in the University Medical Center's Human Electrophysiology Laboratory.

The results of our qEEG work as outlined above were reported to the Acoustical Society meeting in Honolulu (Lauter 1988a).

MRI/MRS. Although within the last few years, anatomical-imaging applications of Nuclear Magnetic Resonance (NMR) have gained wide acceptance as tools in clinical medicine, the physiological-imaging capabilities of NMR are just beginning to be developed. During the summer of 1988, we attempted a pilot study using a combination of Magnetic Resonance anatomical Imaging (MRI) and Magnetic Resonance Spectroscopy (MRS) to explore the potential of these applications of NMR technology for noninvasive study of human brains during rest and activation. Problems related to stimulus delivery and time constraints led us to choose a simple motor activation paradigm, similar to those used previously with qEEG (see above) and PET (more below).

Tests were conducted in the MRI Laboratory at the Waisman Center in Madison WI, using methods developed by William Perman and John Sandstrom; Dr. Perman acted as the experimental subject. Conditions tested during separate 10-min scans were:

1) baseline control: no movement, subject resting quietly with
   eyes closed
2) right-hand flexion at approximately 60/min; eyes closed
3) control condition: subject quiet
4) second control condition: subject quiet
4) right-hand flexion at 60/min.

Testing was conducted using the General Electric 1.5T SIGNA MR scanner in the Magnetic Resonance Imaging Laboratory of the Waisman Center in Madison WI. The

subject reclined with eyes closed on the testing table, with a light cover for warmth. A 20 cm x 15 cm plastic bag filled with acetone and water, designed to serve as an external phantom to assist in normalizing brain signals with regard to the drift of the scanner, was placed on his forehead. Then the table was rolled into the scanner, the room darkened, and the magnetically-shielded door closed.

Data-collection methods for both the initial MRI anatomical series and the subsequent MRS scans taken during each test condition were modelled on those described in detail in Sandstrom et al (Submitted) and Partington et al (Submitted).

Data for the MRI anatomical scan were collected using the standard 17 cm extremity rf coil, and partial saturation spin echo (PS) pulse sequence with TR = 600 ms, TE = 55 ms, 128 phase encoding steps (1.8 mm spatial resolution), 256 readout points (0.9 mm spatial resolution), and two rf excitations. Once the MRI scan was completed, yielding a series of horizontal slices of the subject's brain, the file was examined to determine which slice would provide the best target for MRS analysis. Slice selection was guided by: 1) published information regarding location of the "hand area" of the motor cortex; 2) our own previous results on hand-movement studies using qEEG (approximate position of electrode locations F7/8, cited above); and 3) our previous results on hand-movement monitored with PET (Fig. 21 panel A: white area of response shown in a horizontal section on the left, and in a coronal section on the right). Based on these guidelines, a slice 10mm rostral to the slice representing the most rostral appearance of the ventricles was selected. Within this slice, a 2 $cm^3$ region was selected as the ROI. In the anterior-posterior dimension, this region was centered on the central sulcus, in order to maximize contribution from both motor and somatosensory primary cortex (cf. Fig. 21 panel B for the MRI slice with the ROI indicated by a white box).

The MRS data were collected using broadband proton spectroscopy enhanced via volume selection and water suppression. Volume selection was accomplished using a volume selective pulse sequence, applying three 90-degrees rf pulses, each one selective on a different gradient axis. Water suppression was effected by substituting a binomial 1-3-3-1 180-degrees rf pulse for the standard 180-degrees rf refocussing pulse, and also applying a narrow band (0.5 ppm) presaturating 90-degrees rf pulse having a sine waveform, centered on the water resonance, preceding the sequence.

Data collections were taken under the five test conditions listed above, each requiring approximately 10 minutes. Postprocessing of the data included: 1) magnitude calculation of the Fourier transformed spin echo data, 2) normalization using the external phantom, and 3) subtraction of the baseline spectrum (condition 1) from the spectra collected during conditions 2 through 4.

Figure 22 illustrates the normalized but unsubtracted spectra (the form of the spectra yielded by step 2 above), showing amplitude as a function of Parts Per Million (PPM). This is a spectral dimension that is independent of magnet magnitude, along which the

resonances of various chemicals are expressed relative to tetramethylsilane (TMS), normally assigned a value of 0 ppm. In this figure, spectra for the five data collections are placed serially, one above the other, along an implied third dimension of time; the test condition is indicated for each.

Identification of the molecules represented by each of the peaks here is difficult; even in the invasive dog experiments described by Perman and colleagues (op cit), in which known substances were introduced directly into the ROI during scanning, ambiguities arise due to interactive resonance properties of target and context. However, a number of molecules represented in this PPM range can be defined, and represent substances involved in neural activity.

The subtracted spectra are displayed in Fig. 23, in this case showing amplitude of difference from baseline as a function of PPM. Panel A shows positive changes, panel B shows negative changes. As in Fig. 22, in each panel, results for each of the four test conditions are displayed above one another along an implied third dimension of time; conditions are indicated. In these subtracted spectra, there appear to be changes systematically related to test condition. The most dramatic positive change (Panel A) is in the region of 0.75 ppm, where a large peak appears during both hand-movement conditions, but is much reduced during controls. For negative changes (Panel B), a peak at approximately 0.5 ppm is prominent in both hand-movement conditions, but small to absent in the controls; and there is also a broad negative-change peak centered at -0.5 ppm, which occurs in both hand-movement conditions, but splinters into individual lower-amplitude peaks in the controls.

The spectral region between approximately 1.0 ppm and -1.0 ppm, which seems to be the focus of change under these conditions, is highlighted in Fig. 24, with positive changes again shown in Panel A and negative changes in Panel B. As mentioned above, exact assignment of molecules to the points of change is difficult. However, two aspects of the results suggest that these spectra are in fact revealing changes directly related to activation. First, spectra for both hand-movement conditions are similar in the patterns of positive and negative changes with regard to baseline, and spectra for these two activation conditions are different from those collected under the two control conditions.

Second, and perhaps most encouragingly, the details of these changes are sensitive to the chronology of the test conditions: specifically, graded residual effects of the changes observed during the first hand-movement condition can be observed in the subsequent control conditions. For example, in the positive changes (Fig. 24, Panel A), note how the 0.75 ppm peak in R1 falls to half this height in C1, half again as small in C2, to return to full R1-like amplitude in R2. The negative change (Panel B) at 0.5 ppm is broad and high for R1, becomes very narrow and smaller in C1, disappears in C2, to return in its original broad, high-amplitude form in R2. Finally, the negative change centered at -0.5 ppm is clearly shrinking from R1 to C1 to C2, but regains its original breadth and amplitude configuration at R2.

These patterns are summarized in Figs. 25 and 26. The three panels of Fig. 25 present a simplified spectral profile representing 10 resonance locations, ranging from 1.00 PPM to -1.00 PPM. Panel A displays a comparison of this spectral profile as observed in the two right-hand-movement conditions. Note that the two profiles shown here are almost identical. Panel B displays a similar comparison, based on data from the two control conditions; the "drift toward zero" is apparent in this comparison, both for spectral locations that were originally negative re baseline (e.g., +0.5 PPM and -0.25 PPM) as well as for locations that were originally positive re bseline (+0.75 PPM and +0.25 PPM). Panel C illustrates the extreme comparison, between the profile for right-hand condition #1 and control condition #2: the flattening of the profile is apparent at all locations except 0.0 PPM which shows virtually no change no matter the condition.

In Figure 26, data for the four spectral locations showing the most dramatic differences from test to control are graphed in a different way, plotting relative amplitude as a function of condition with spectral location as a parameter. The positive change in the resonance at +0.75 PPM (open squares) seen with hand movement during condition R1 falls toward zero across the two controls, then returns to high amplitude for R2. The other resonances shown (closed symbols) show negative changes from baseline with hand movement, which are reduced with C1 and go to zero at C2, to return to negative values with condition R2.

This pattern of residual activity observed after the cessation of activation is reminiscent of that observed in our other research using noninvasive methods for monitoring human brain function. For example, PET images from a particular 40-sec scan often reflect not only the condition under test during that scan, but also conditions tested on previous scans. Figure 27 Panel A shows blood-flow images from an experiment on hand movement, contrasting images taken under conditions of no movement (control: shown in upper L) with images for left-hand flexion (upper R) and right-hand flexion (lower L).

These images show not only responses to current activation (the white areas indicate regions of high blood flow), but also "residual" activity left from previous scans. The actual test order of these scans was: left-hand movement (the only area of activation is in the contralateral, R-hemisphere hand area) followed by right-hand movement (L-hemisphere white spot, with "residual" activity in the R hemisphere from the previous left-hand condition) followed by control ("residual" activation left from both the unilateral movement conditions). Note that the time course during which the residual effects were observed in this PET study is similar to that for MRS: a minimum of 15 minutes separated each of the PET scans, such that the control scan shows the effects of an activation condition (left-hand movement) completed at least one-half hour earlier.

A similar instance is shown in Fig. 27 Panel B, summarizing data from our qEEG experiments, for one subject's beta power asymmetries comparing electrode locations T3 and T4 under four conditions: dichotic syllables with attention to the right ear ("S"),

dichotic tone patterns with attention to the left ear ("T"), and three controls ("C1" collected at the beginning, "C2" collected after syllables, and "C3" collected after tones). Note that the asymmetries during control scans seem to be affected by preceding activation conditions, shifting from the baseline value either 1) in the direction of the syllable asymmetry, following syllables, or 2) in the direction of the tone asymmetry, following tones. This "residual" effect of previous test conditions on qEEG data collected under control conditions seems exactly parallel to the observations in the MRS data. In both cases, patterns of brain activity monitored during a control condition following an activation condition share some characteristics of patterns observed during the activation condition. For MRS, where a second control was collected, we were able to observe a second data point in the time course of the "decay" of the effect of activation.

Since MRS subtracted spectra such as those shown in Fig. 24 are derived by comparison of the actual spectrum for each condition with the original baseline, it is to be expected that eventually a control condition should result in a flat line—i.e., no difference when compared with baseline. Although this might never actually occur, since the "baseline" might represent only one characteristic resting measurement, there are indications that the two control subtracted spectra shown in Fig. 24 are moving in that direction, particularly for the plot of Control 2's negative changes (Panel B), where more than half of the second control's subtracted spectrum is at zero. We do not yet have qEEG or PET data to suggest how long after stimulation the original resting power asymmetry is restored, but there are indications that initial resting values, such as the 34% RHA shown for "C1" of Fig. 27 panel B, is characteristic of an individual subject, and can be replicated over sessions.

The MRS experiment described here is admittedly a preliminary one, since only those spectral changes occurring in one ROI tested under a single activation condition in one subject were observed. However, we believe that the results are quite encouraging, particularly in view of their resemblances with data collected using other noninvasive methods monitoring human brains under similar test conditions.

Results of the experiment described here have been detailed in a privately-circulated report (Lauter 1988b).

PET. Before the PI left St. Louis in spring 1988, an agreement was reached regarding future interactions between the PI and the PET lab of the Washington University Mallinckrodt Institute of Radiology in St. Louis: 1) if the PI could obtain the appropriate hardware for a "data analysis satellite system" to be located in Tucson, the Mallinckrodt lab would provide copies of all PET data files collected by the PI in St. Louis during 1981-1985, together with appropriate image-processing software for analyzing the data; and 2) if analysis of these results proved interesting, the PI could propose future projects, with data collection to be conducted in St. Louis, and data analysis done on the Tucson satellite system.

On the strength of this informal arrangement, the PI submitted an equipment proposal to the DoD-URIP for purchase of a minicomputer-based data analysis satellite system, following specifications provided by the St. Louis laboratory. Upon funding of this proposal (AFOSR 87-0003), the hardware, consisting of a Perkin-Elmer 3205 minicomputer, a Ramtek MC68000 Color Display Controller with 19" RGB monitor, and a Matrix 3000 film recorder, was purchased, and installed during spring of 1987. Negotiations with the St. Louis laboratory continued until April 1988, when all data files and relevant image-processing software were delivered to our PET data-analysis lab in Tucson.

As outlined above, the PET data analysis laboratory is housed in a large basement laboratory room in the Psychology Building on the UA campus. It has been used to illustrate our PET work to visitors, and as an effective visual aid in guest lectures made by the PI to both Speech & Hearing and Psychology department courses.

As a result of the variety of test conditions included, the data library can support a number of research projects. The first of these is in progress, and involves establishing a background for future studies of asymmetries in regional cerebral blood flow (rCBF) in normal subjects. We are interested in rCBF evidence of three types of brain asymmetries: 1) resting asymmetries (such as those documented in our qEEG pilot experiments--see above); 2) asymmetries based on "side of space," e.g., for the auditory system, differences in activation based on side of input; and 3) asymmetries based on physical aspects of stimuli, such as those suggested in our dichotic-listening results.

The project currently in progress in the PET laboratory addresses the question of resting asymmetries, with the goal to compile baseline measures of hemispheric differences in rCBF during "resting" conditions, for later comparison of rest conditions across subjects, and comparison of rest with stimulation conditions within subjects. Note that because of some repeated sessions in the original library, baseline comparisons can also be made within subjects across sessions, to examine whether a subject's pattern of resting asymmetries changes with time.

To date, resting rCBF has been measured for a total of six subjects (three female, three male), for each hemisphere in each of six slices (the more rostral six of the seven available) for each of two "control" scans per session per subject. As illustrated in Fig. 28, Panel A, data-analysis software provided by the St. Louis PET laboratory made it possible to: 1) display each slice to be measured; 2) obtain a calculation of the midline of the slice (based on midline of the "field of view" containing that slice); and 3) using a cursor under observer control (the "irregular regions" routine included in the programs), outline each hemisphere, and obtain a calculation of both the number of pixels and the mean rCBF within the outlined region. After the values for each hemisphere in the six slices of a subject's rest scan were provided by the program, they were summed across slices for each hemisphere, and a right/left quotient expressing asymmetry in the brain as a whole calculated both for number of pixels and rCBF.

Results on rCBF asymmetries for the six subjects completed to date are presented in Fig. 28 Panel B, with slice-by-slice rCBF asymmetries compared for the first control scan in each session (open bars) and the last control scan (striped bars). The six graphs shown here indicate that: 1) all six subjects show a slightly higher average resting rCBF in the right hemisphere than in the left (this is in agreement with earlier published reports of rCBF resting asymmetries); 2) there are between-subject differences in the pattern of slice-by-slice asymmetry; and 3) there is good within-subject consistency in slice-by-slice asymmetry.

These results on rCBF in the two hemispheres are complemented by measurements taken during the same image analysis related to size of the two halves of each displayed slice, expressed as number of image pixels per side. For three of the six subjects, there were more pixels in the summed six slices of the right hemisphere than of the left; this is in agreement with MRI anatomical scans of normal subjects, most of whom show a right hemisphere slightly larger than the left. However, two of the six subjects showed larger left than right hemispheres, and one subject showed equal number of pixels on the two sides.

Measurement reliability for both rCBF and pixel asymmetry analysis was calculated using Pearson's product moment correlation for 4% of all measures (selected at random): $r = 0.99$ ($p < .01$) for rCBF, and $r = 0.99$ ($p < .01$) for the number of pixels. Pearson's r was also used to examine the within-subject consistency of the right/left asymmetry quotient between the two resting scans from each subject's session: $r = 0.99$ ($p < .01$) for rCBF, and $r = 0.86$ ($p < .05$) for number of pixels measured in each hemisphere.

Analysis of the other control scans in our data library is continuing, with plans for later stages of this project to examine global asymmetries during all of the activation conditions represented. This survey of global brain asymmetries during control and activation will serve as a background for measures to be taken in the CNS Project, where we will be focusing on cerebral asymmetries in response to three test sounds. We will also make use of the current PET-based observations on brain size asymmetries in the within-individual comparisons planned for the CNS Project subjects between hemisphere size differences documented with PET and with MRI.

The first report on tonotopic organization in human auditory cortex studied with PET was published early in the period (Lauter et al 1985b). A new summary of our PET findings with auditory stimulation was given as a meeting presentation (Lauter et al 1987) and subsequently expanded into a publication (Lauter et al 1988). Preliminary results of the resting global rCBF results will be reported to the Acoustical Society of America in spring 1989 (Lauter & Plante 1989).

## SIGNIFICANCE AND EXTENSION TO NEW PROJECT

All of these projects represent both continuations of previous interests, and preparations for future research such as that planned for the Coordinated Noninvasive Studies (CNS) Project (AFOSR 88-0352). All of the established laboratory facilities will serve the new project, and data collected in each of the research areas will serve as a basis for design and interpretation of related new results.

Each subject in the CNS Project will first be tested behaviorally to determine: 1) sidedness (using the Harris 1974) questionnaire; 2) hearing acuity for each ear; and 3) ear advantages for three sets of sounds, 200-ms pure-tone patterns, 50-ms pure-tone patterns, and synthetic stop CV syllables. For the dichotic battery, sound preparation and testing guidelines will be based on procedures established in the previous period.

Second, each subject will receive an MRI anatomical scan, collected on a fee-for-service basis at the University Medical Center's MRI Center in Tucson. These scans will be examined using procedures developed during the previous period by a graduate-student collaborator, E. Plante, for asymmetries in cortical and subcortical regions related to auditory processing.

Next, each subject will be scheduled for a qEEG session, to be tested with each of the dichotic sound sets with alternating right- and left-ear attention. Each session will begin with a control scan, followed by three hand-movement conditions, to provide cerebral-asymmetry "anchor points;" test segments on each of the sound sets will be separated by control scans with no stimulation; and the session will conclude with a final control scan. All procedures, analysis techniques, and interpretation of these results will be guided by the pilot work on qEEG described above.

The last two tests for each CNS Project subject will be: 1) a PET scan, followed by 2) brain monitoring using MEG. Facilities for conducting these two test series are currently being sought by the PI, with negotiations initiated and continuing at this time. Both types of test will include separate scans during non-stimulation control conditions as well as during stimulation with each of the three sound sets. It is expected that the PET-scan results for each subject will help localize any discrete brain responses to the three sound sets, including subcortical responses such as those we have observed in previous subjects (cf. Lauter et al 1985a), and thus aid in locating regions of interest peculiar to each subject, ensuring maximum efficiency in MEG testing. The PET experiments will build on our PET experience and results detailed above, and the MEG testing will be guided by findings in an abbreviated pilot experiment on a 7-channel MEG, contrasting responses in one subject to pulsed pure tones at two frequencies and a syllable sequence, not described here.

In summary, the three years of AFOSR 85-0379 have provided the opportunity to establish a strong base, both in terms of laboratory facilities as well as procedural experience and guidelines, supporting our research projects planned for the immediate future.

REFERENCES

Baldwin, C.M., Lauter, J.L., and Facciola, P.C. (1988a)  Gender differences in the production of vocal emotional expressions, to Western Psychological Association, San Francisco CA.

Baldwin, C.M., Lauter, J.L., and Facciola, P.C. (1988b)  Toward a taxonomy of vocal expressions of emotion, to Society for Psychotherapy Research, Santa Fe NM.

Baldwin, C. & Lauter, J.L. (In preparation) Simulated emotional expressions by actors and non-actors: some general acoustical correlates.

Berlin, C.I.,  Hood, L.J., and Allen, P. (1984)  Asymmetries in evoked potentials. In C.I. Berlin (Ed.), Hearing science. College-Hill: San Diego;  pp. 461-478.

Harris, A.J. (1974)  Harris Tests of Lateral Dominance, 3rd ed. The Psychological Corp: NY.

Lauter, J.L. (1982) Dichotic identification of complex sounds: absolute and relative ear advantages. J Acoust Soc Amer 71: 701-707.

Lauter, J.L. (1983) Stimulus characteristics and relative ear advantages: a new look at old data. J Acoust Soc Amer 74: 1-17.

Lauter, J.L. (1984) Contralateral interference and ear advantages for identification of three-element patterns. Brain and Cognition 3: 259-280.

Lauter, J.L. (1988a)  Functional-activation asymmetries in normal humans studied with quantitative EEG (qEEG):  first tests in the CNS Project, to Acoustical Society of America, Honolulu HI.  Abstract: J Acoust Soc Amer 84: S57.

Lauter, J.L. (1988b) Functional brain activation in normal humans studied with volume-selective water-suppressed in-vivo NMR proton spectroscopy: a preliminary report. (Unpublished MS)

Lauter, J.L. & Baldwin, C. (In preparation)  Acoustical correlates of emotional expression. I. Temporal characteristics of suprasegmental and segmental aspects of simulated emotion in semantically relevant as well as "neutral" sentences.

Lauter, J.L. & Hirsh, I.J. (1985) Speech as temporal pattern: a psychoacoustical profile. Speech Communication 4: 41-54.

Lauter, J.L. & Karzon, R.G. (1987)  Individual differences in auditory evoked potentials: middle-latency responses, to Acoustical Society of America, Indianapolis. Abstract: J Acoust Soc Amer 81: S8.

Lauter, J.L. & Karzon, R.G. (In preparation, a)  Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability.  III.  Influence of individual-subject characteristics on group comparisons of ABR latency and amplitude variability.

Lauter JL & Karzon, R.G. (In preparation, b)  Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability. IV.  Consistency of responses representing three levels of the auditory CNS:  ABRs, middle-latency, and cortical EPs.

Lauter, J.L. & Loomis, R.L. (1985)  Individual differences in auditory evoked responses: comparisons of between-subject and within-subject variability in brainstem and cortical waveforms, to Acoustical Society of America, Austin. Abstract: J Acoust Soc Amer 77: S65.

Lauter, J.L. & Loomis, R.L. (1986a)  Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability.  I.  Absolute latencies of brainstem vertex-positive peaks.  Scand Audiol 15: 167-172.

Lauter, J.L. & Loomis, R.L. (1986b)  Individual differences in evoked responses: second report, to Acoustical Society of America, Cleveland. Abstract: J Acoust Soc Amer 79: S5.

Lauter, J.L. & Loomis, R.L. (1988)  Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability.  II.  Amplitude of brainstem vertex-positive peaks.  Scand Audiol 17: 87-92.

Lauter, J.L. & Lu, F.-L. (1987a)  VOT variability in stop-consonant productions by bilingual speakers of English and Mandarin Chinese, to Acoustical Society of America, Miami. Abstract: J Acoust Soc Amer 82: S115.

Lauter, J.L., & Lu, F.-L. (1987b)  VOT variability in Mandarin Chinese: interactions with tone, to International Society for Phonetic Sciences, Miami Beach.

Lauter, J.L. & Lu, F.-L. (In preparation)  VOT variability in Mandarin Chinese: interactions with tone in Chinese and English productions.

Lauter, J.L. & Pearl, N. B. (1986)  VOT variability: Within-subject and between-subject measurements of stop-consonant production by female talkers of English, Japanese, Navajo, and Spanish, to Acoustical Society of America, Anaheim. Abstract: J Acoust Soc Amer 80: S62.

Lauter, J.L. & E. Plante (1989)  Global brain asymmetries in regional cerebral blood flow (rCBF) observed during resting conditions with positron emission tomography (PET): establishing a baseline for experiments on brain asymmetries and complex sounds in the CNS Project.  To be presented to Acoustical Society of America, spring 1989.

Lauter, J.L., Herscovitch, P., and Raichle, M.E. (1983a) Human auditory cortex: a preliminary report on studies using positron emission tomography (PET), to Acoustical Society of America, Cincinnati. Abstract: J Acoust Soc Amer 73: S60.

Lauter, J.L., Herscovitch, P., Formby, C., and Raichle, M.E. (1983b)  Tonotopic organization in human auditory cortex as revealed by regional changes in cerebral blood flow,  to XI International Symposium on Cerebral Blood Flow and Metabolism, Paris. Abstract: J Cereb Blood Flow Metab 3: S248-249.

Lauter, J.L., Herscovitch, P., Formby, C., and Raichle, M.E. (1984)  Changes in human regional cerebral blood flow in response to pure tones, to Acoustical Society of America, Norfolk. Abstract: J Acoust Soc Amer 75: S14.

Lauter, J.L., Herscovitch, P., and Raichle, M.E. (1985a)  Human auditory central nervous system: visualization of cortical and subcortical centers with regional cerebral blood flow measurements on the PETT-VI, to Acoustical Society of America, Austin. Abstract: J Acoust Soc Amer 77: S94.

Lauter, J.L., Herscovitch, P., Formby, C., Raichle, M.E. (1985b) Tonotopic organization in human auditory cortex revealed by positron emission tomography. Hearing Research 20: 199-205.

Lauter, J.L., Herscovitch, P., and Raichle, M.E. (1987) Human auditory physiology studied with positron emission tomography, to Auditory Pathway: Structure and Function conference, Prague.

Lauter, J.L., Herscovitch, P., and Raichle, M.E. (1988) Human auditory physiology studied with positron emission tomography. In J. Syka and R.B. Masterton (Eds.), Auditory Pathway. Plenum: NY; pp. 313-317.

Lauter, J.L., Solomon, N.P., and Baldwin, C. (Submitted) VOT variability in repeated productions of nonsense stop CV syllables. (To: Journal of the Acoustical Society of America).

Lord-Maes, J.M. & Lauter, J.L. (1989) Comparisons of between- and within-subject variability in repeated-measures auditory brainstem responses (ABRs) in 5-6-year-old children. To be presented to Acoustical Society of America, spring 1989.

Oyler, R. & Lauter, J.L. (1989) Comparisons of between- and within- subject variability observed in repeated-measures auditory brainstem responses (ABRs) in healthy 10-year-old males. To be presented to Acoustical Society of America, fall 1989.

Partington, C.R., Sandstrom, J.C., and Perman, W.H. (Submitted) Volume localized in vivo spectroscopy of canine status epilepticus.

Sandstrom, J.C., Perman, W.H., and Partington, C.R. (Submitted) A volume selective pulse sequence with water suppression for in vivo proton spectroscopy

(voice onset)

#1  #4

#2  #5

#3  #6
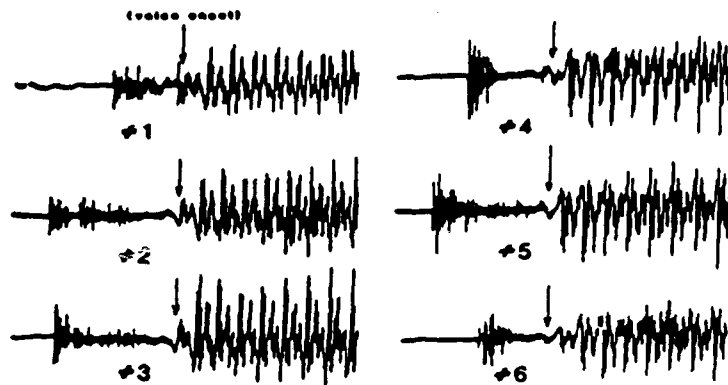
Figure 1. Waveforms of the VOT portion selected from six tokens of "ha-ga" produced by a female native English speaker. Arrows mark onset of voicing. Such displays were used to measure VOTs in the study being described.

**6 repetitions of 6 stops (Amer. Eng.)**

VOT (ms)

80

60

40

20

0

1    2    3    4    5    6    X̄

repetition

b
d
g
p
t
k
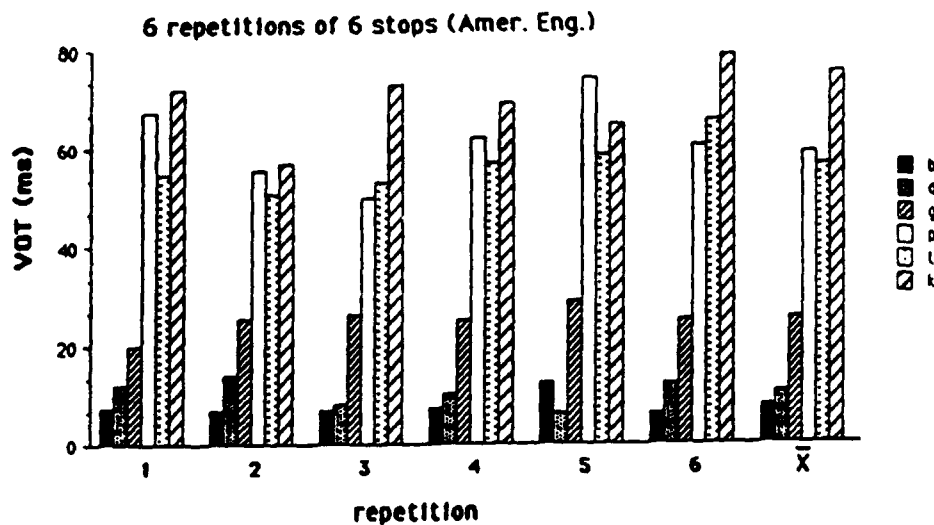
Figure 2. Bargraph illustrating absolute VOT values for each of six stop consonants produced in six repetitions in a "ha-Ca" context by one female native English speaker. Mean VOT for the six phonemes is shown at the right. Note that within individual repetitions, the distinction between voiced/voiceless is maintained, but there are ambiguities in VOT timing within each category.
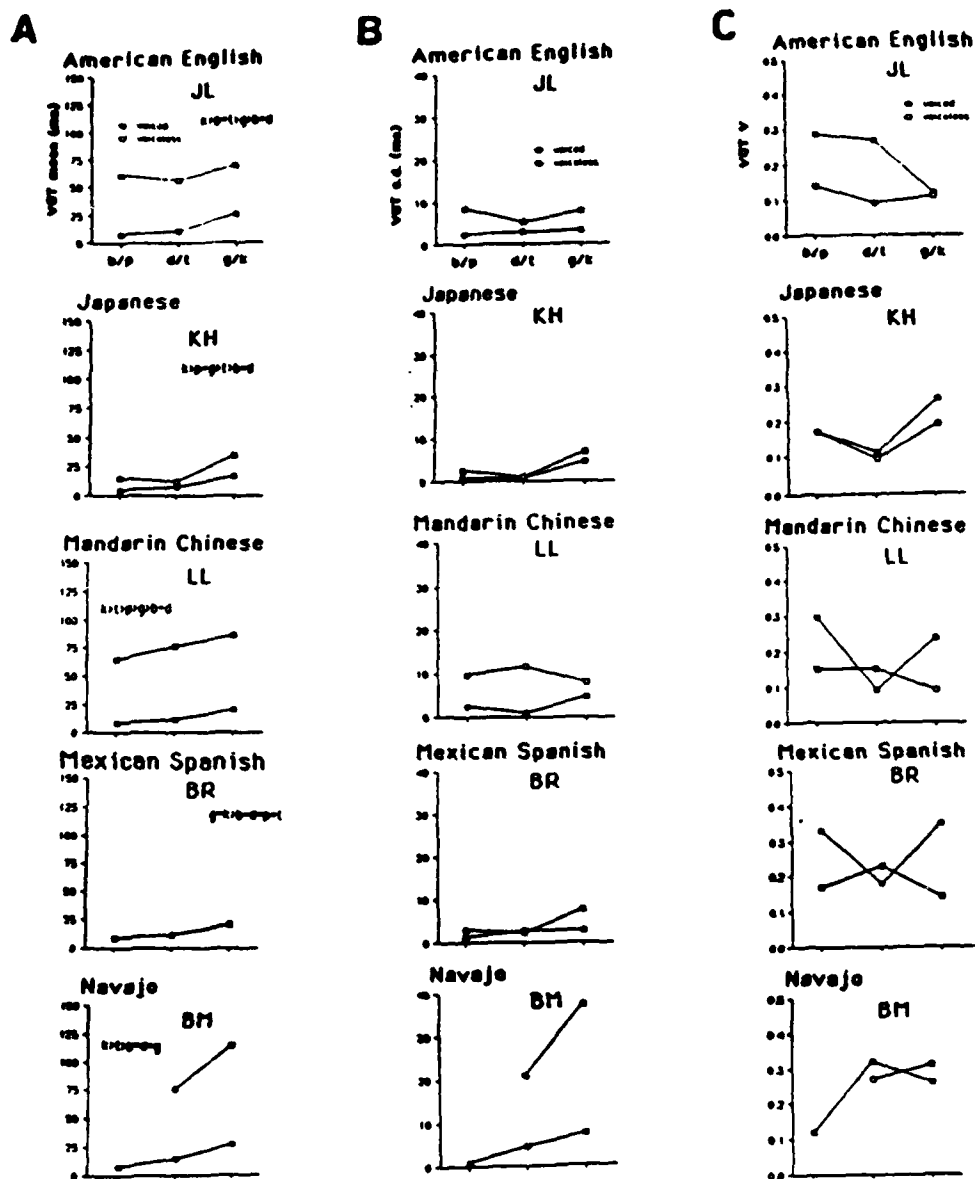
Figure 3. Panel A: VOT mean in ms as a function of stop consonant, for one female native speaker of each of five languages. Insets show results of Neuman-Keuls test of effects demonstrated in an analysis of variance, indicating significant (p<.05) differences in mean VOT among the phonemes. Note the language-specific patterns for voiced/voiceless distinctions: yes for English, Mandarin, and Navajo; no for Japanese and Mexican Spanish. Panel B: VOT standard deviation in ms for the productions represented by means in Panel A. Panel C: VOT variability (V = s.d./X) for the same productions. These patterns may be language-universal (note the similarity between the Mandarin Chinese and Mexican-Spanish talkers' patterns).
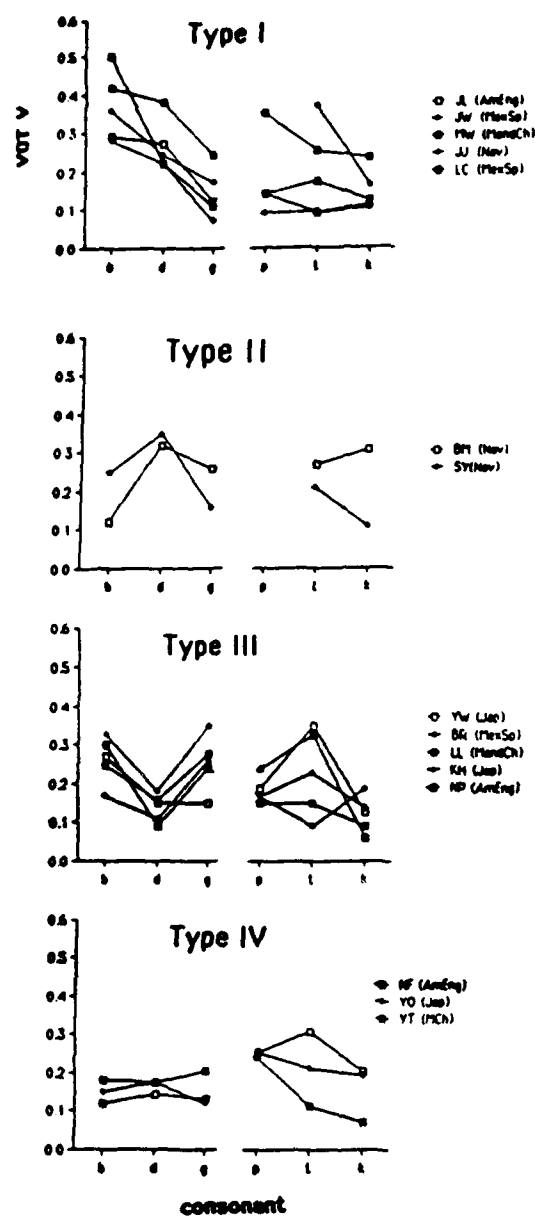
Figure 4. VOT variability patterns observed for 15 adults tested to date. Note the mixture of languages represented in each pattern, suggesting the universality of these pattern categories.

# FO contour: simplified shape
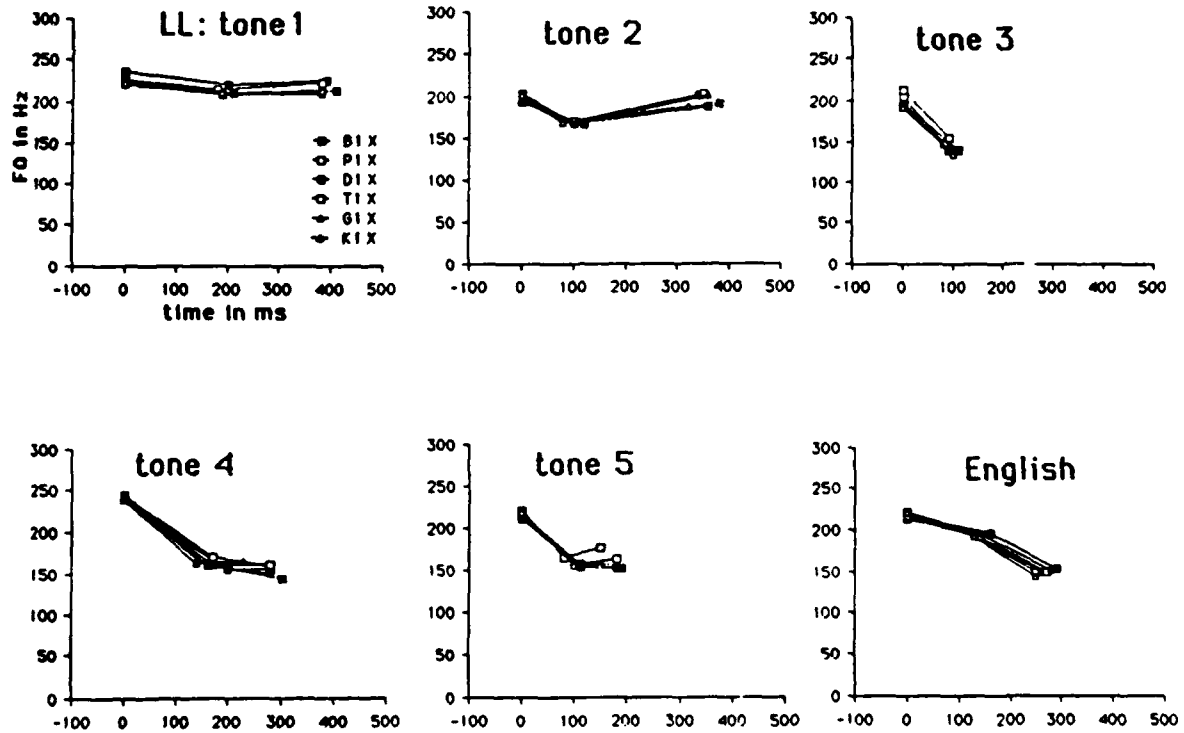


Figure 5. Schematics of fundamental frequency contours used by one female native speaker of Mandarin Chinese to produce nonsense "ha-Ca" tokens on each of the five Mandarin tones and for similar productions of English stop consonants.

# A

## VOT variability (V)
### (English vs. Mandarin 4)



# B

## FO contour: initial frequency and total duration
### (5 Mandarin tones + English)
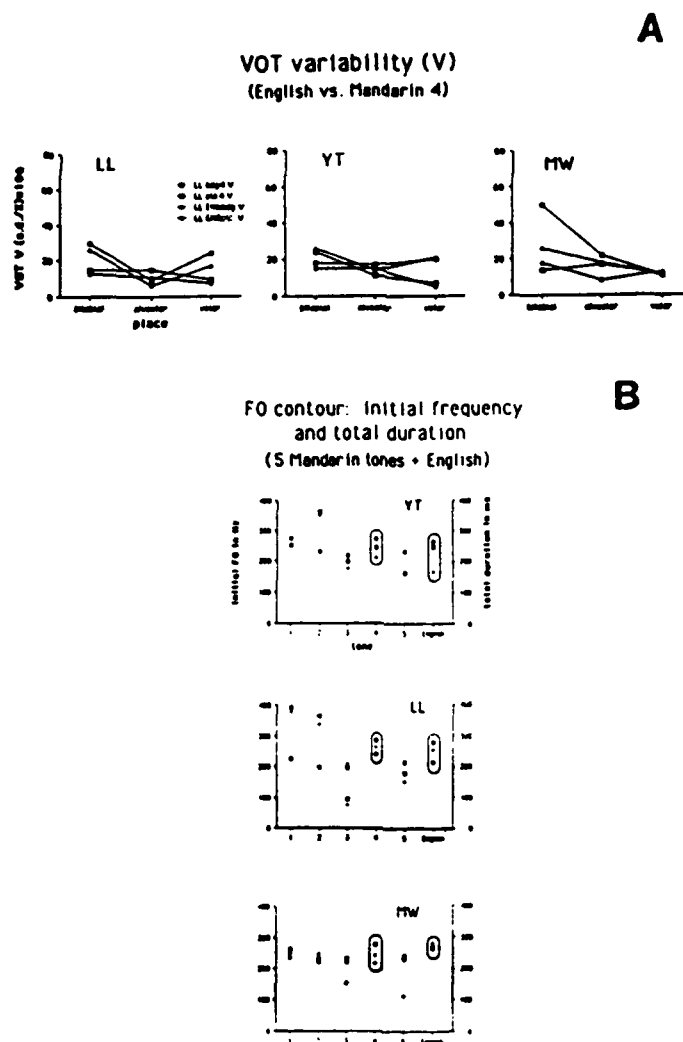


Figure 6. Panel A: VOT variability patterns shown by three female native speakers of Mandarin Chinese for stop-consonant VOTs produced on Mandarin tone 4 and as English phonemes. Panel B: comparison of two characteristics of the tone contour used by the three female speakers for their nonsense productions: for all three speakers, the match of values for initial FO and total duration is best for the Mandarin tone-4 and the English tokens.

"Of course I love you" (NEUTRAL)

"Of course I love you" (HAPPY)

"Of course I love you" (SAD)
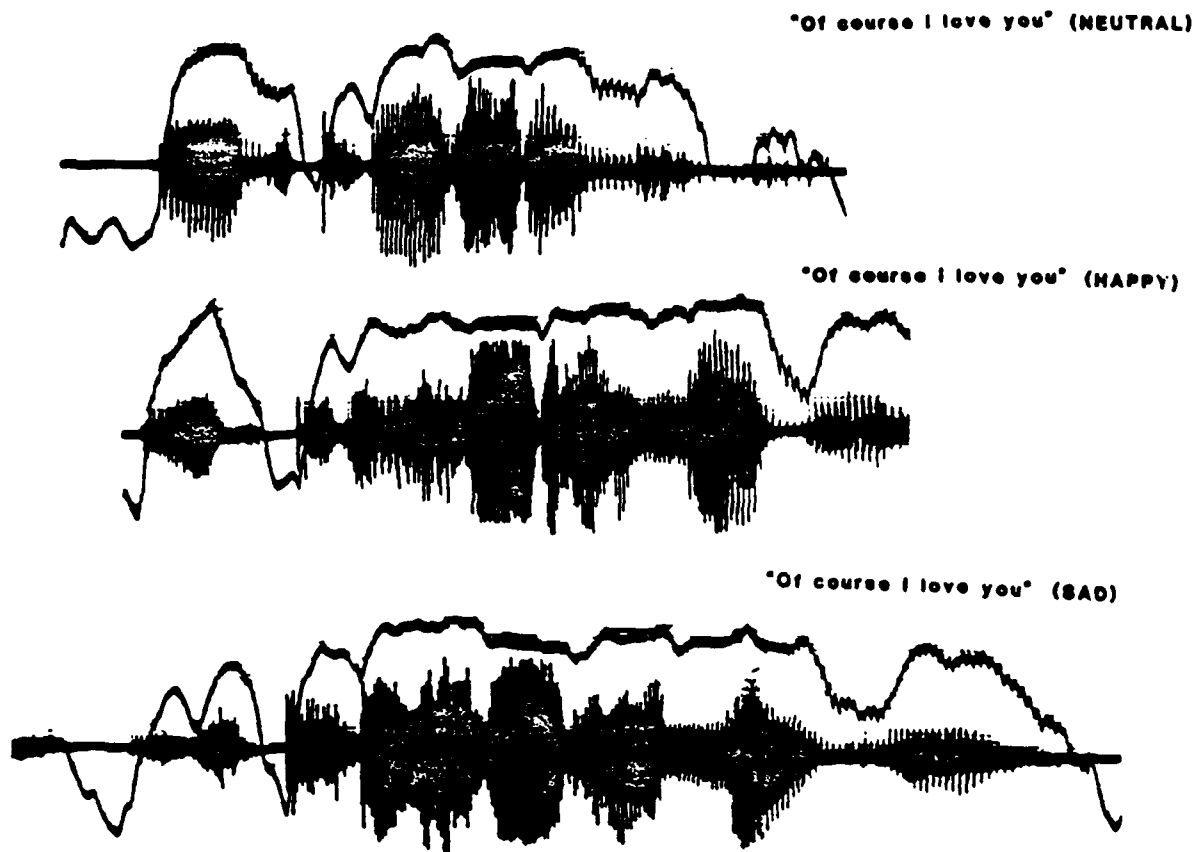
Figure 7. Waveforms and superimposed amplitude contours of the sentence "Of course I love you," spoken by a male actor with three tones of voice: neutral, happy, and sad. These graphs illustrate the changes in both suprasegmental (e.g., overall duration) and segmental (e.g., amplitude of /k/ burst) speech characteristics which accompany changes in emotional expression.

**A**

**"amplitude profiles" (S10)**

EMOTIONS

- □ happy
- ▲ sad
- ■ angry
- ◆ surprise
- ● fear
- ✕ disgust

NEUTRALS

**B**

**overall duration (S10)**

- □ neut dur
- ▨ emot dur
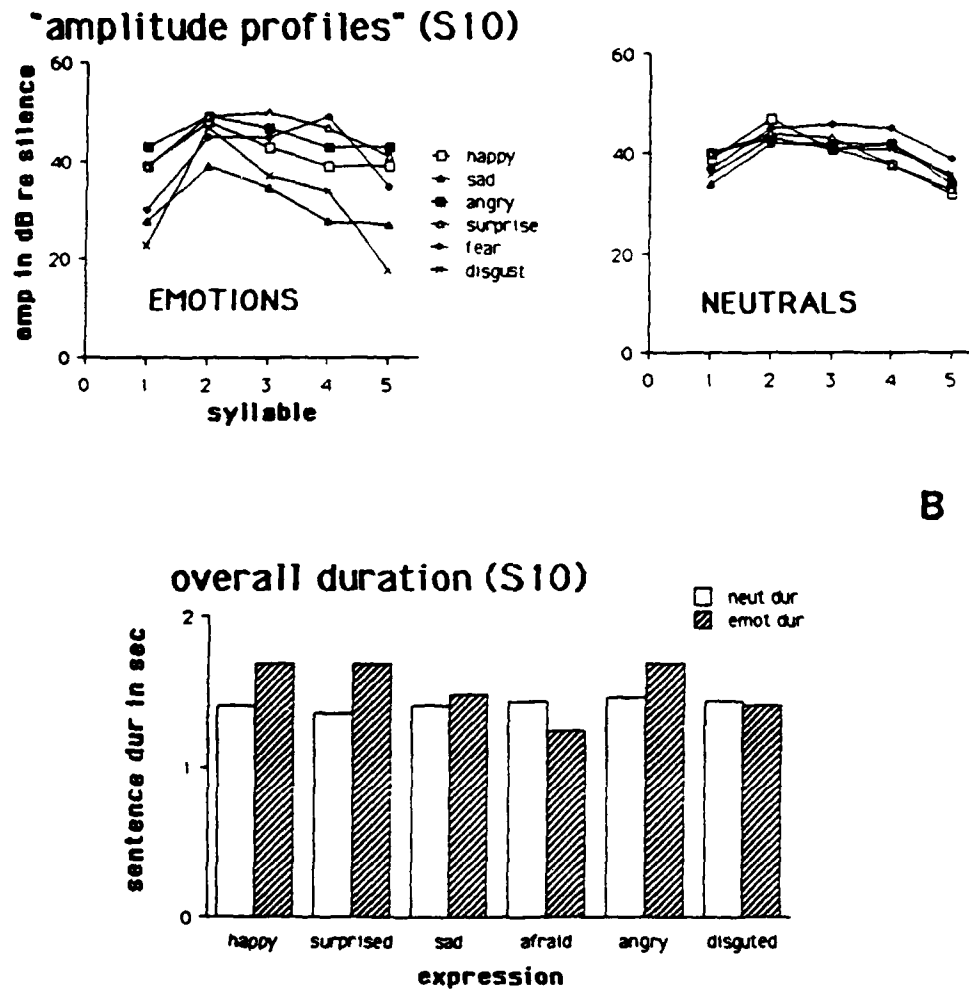
Figure 8. Sample of results of quantitative analysis of waveforms such as those in Fig. 7. Panel A: Comparison of pattern of syllable amplitude peaks for sentences spoken by one subject with different emotional colorations (graph on the left) vs. amplitude profiles used in six instances of the same sentence spoken with a neutral expression. Panel B: Overall duration of the same sentences analyzed in Panel A: note the consistency of overall duration for the neutral sentences, and the fluctuations in duration with emotion.
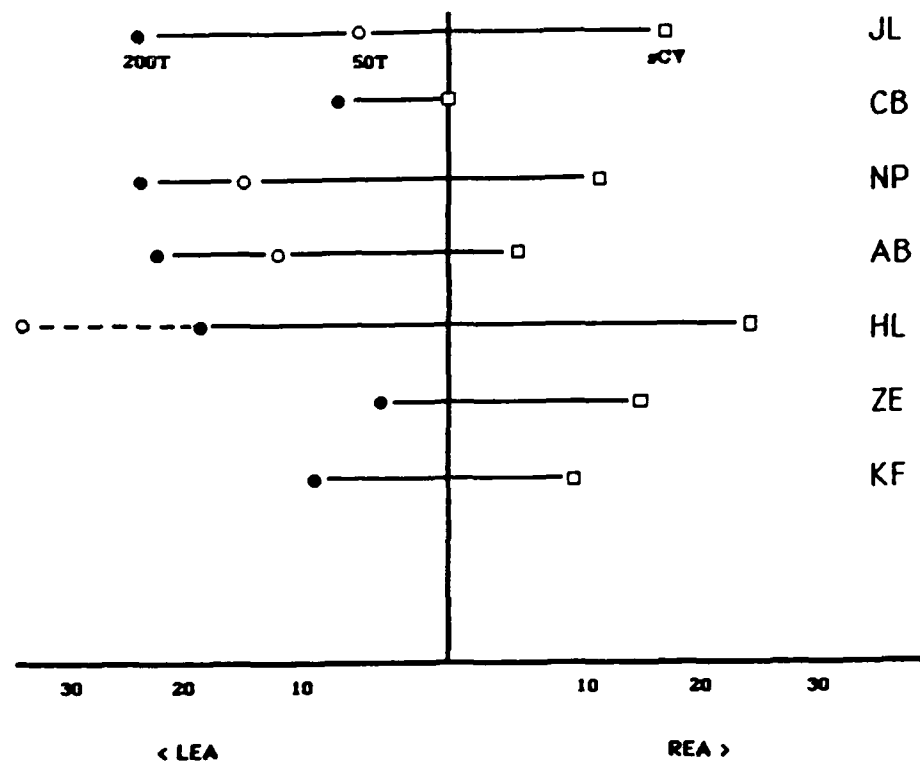
Figure 9. Relative ear advantage plot showing results of dichotic testing with seven listeners: the top five are English speaking, ZE is a Guatemalan-Spanish female, and KF is a Japanese male. Ear differences are shown for three sound sets: pure-tone patterns made with three tones within one critical band centered at 1480 Hz, with 200 ms between tone onsets ("200T" indicated with filled circles), similar patterns with 50 ms between tone onsets ("50T" indicated with open circles), and synthetic stop-CV tokens with English VOT values ("sCV" indicated by open squares). For all listeners, the slow tone patterns evoked EAs to-the-left of those for the syllables. For those listeners tested on the 50T sounds, EAs for these faster tone patterns were intermediate between those for the slower patterns and those for the syllables.

A

## ABR latency
### (7 Ss over 8 sessions)



## ABR latency stability



B

## Replication of ABR latency
### stability profiles

◆ 1st 4 sessions
◆ 2nd 4 sessions
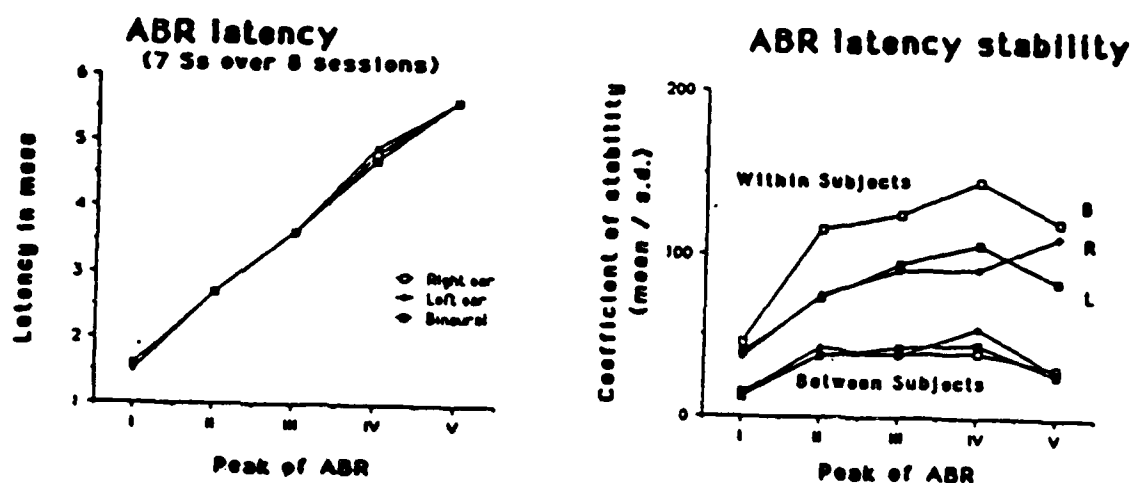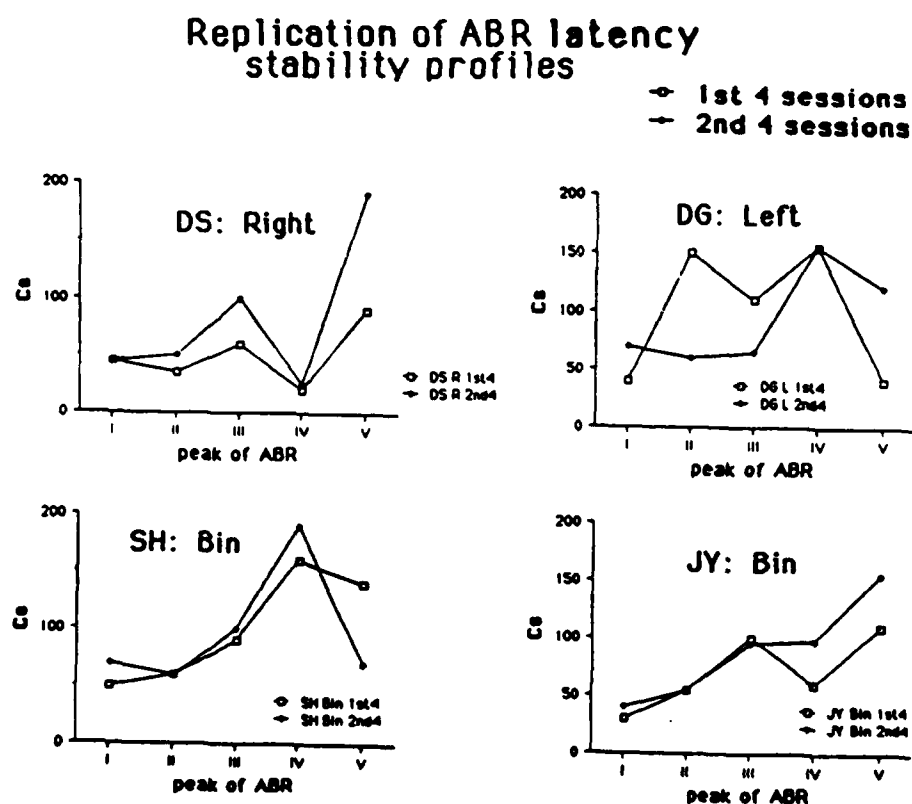


Figure 10. Panel A: Comparison of absolute latency of auditory brainstem response (ABR) waveform peaks (shown on the left) and latency stability (shown on the right). The latency stability measures show distinctions based on peak, group comparisons (i.e., between- vs. within-subject consistency) and ear of stimulation. Panel B: ABR latency variability profiles for four individual subjects, illustrating good replicabilty of these profiles from the first four weeks of testing (open symbols) to the second four weeks (filled symbols).

**Figure 11.** Panel A: Comparison of absolute amplitude of auditory brainstem response (ABR) waveform peaks (shown on the left) and amplitude stability (shown on the right). The amplitude stability measures show distinctions based on peak, group comparisons (i.e., between- vs. within-subject consistency) and ear of stimulation. Panel B: ABR amplitude stability profiles for five individual subjects, on different measures, illustrating good replicability of these profiles from the first four weeks of testing (open symbols) to the second four weeks (closed symbols). There is also an instance of one subject's profile coming to resemble another subject's profile (cf. KP's (B-R) profile which takes two months to assume the shape that SA's (B-R) profile has from the beginning).

# Amplitude variability
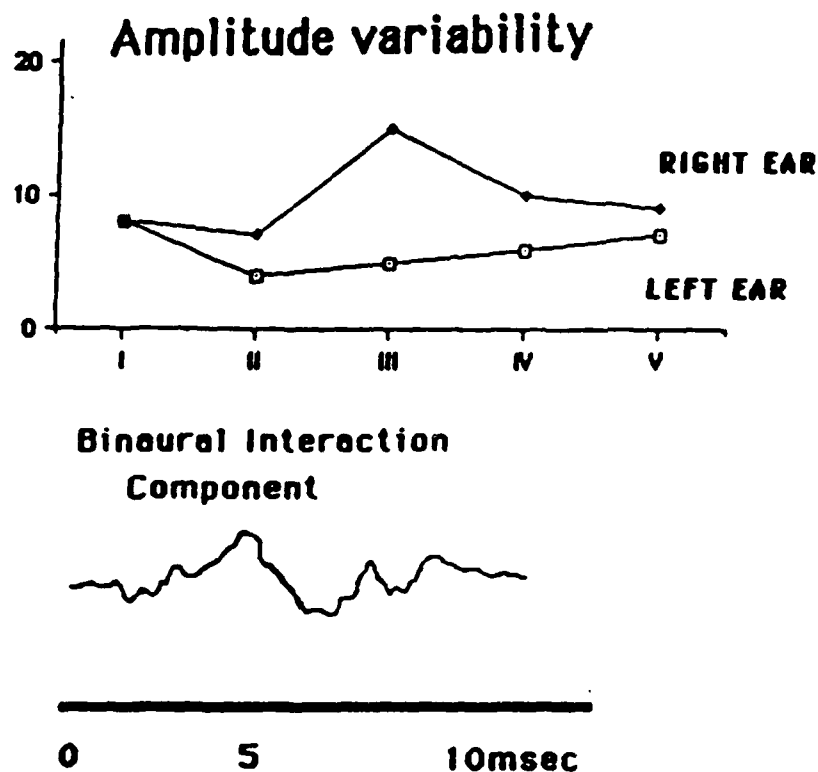


Figure 12. Illustration of two approaches to identifying ear differences in ABR waveforms: top graph is an amplitude stability profile for one subject (cf. Fig. 11), showing higher consistency in the right-ear vs. left-ear response at peak III; lower graph is a tracing of an ABR derived waveform, obtained by adding ipsilateral and contralateral-referenced waveforms of left- and right-ear responses, and then subtracting the two summed waveforms—the peak around 5 ms is designated the "binaural interaction component," and may be quantified in terms of both direction and magnitude to indicate ear difference (cf. Berlin et al 1984).

A
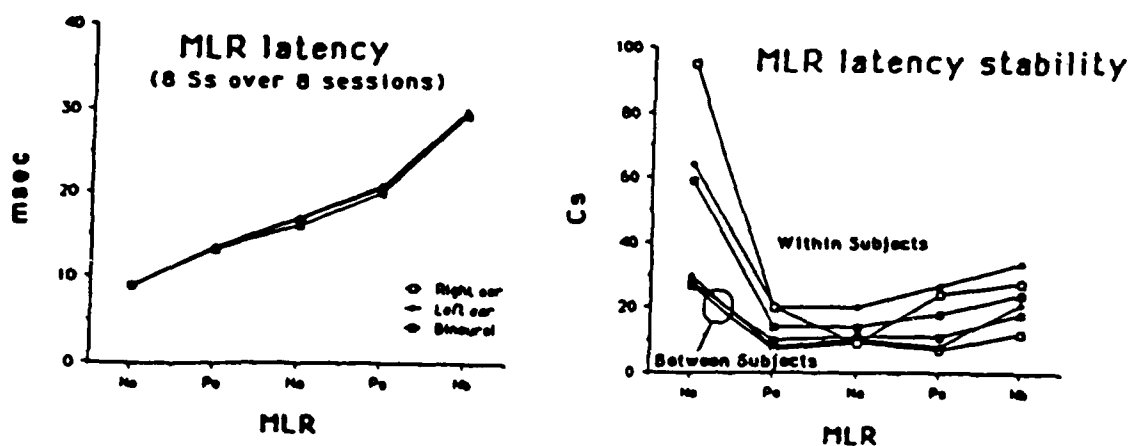


B

## Replication of MLR stability profiles:
## 1st 4 sessions compared with 2nd 4
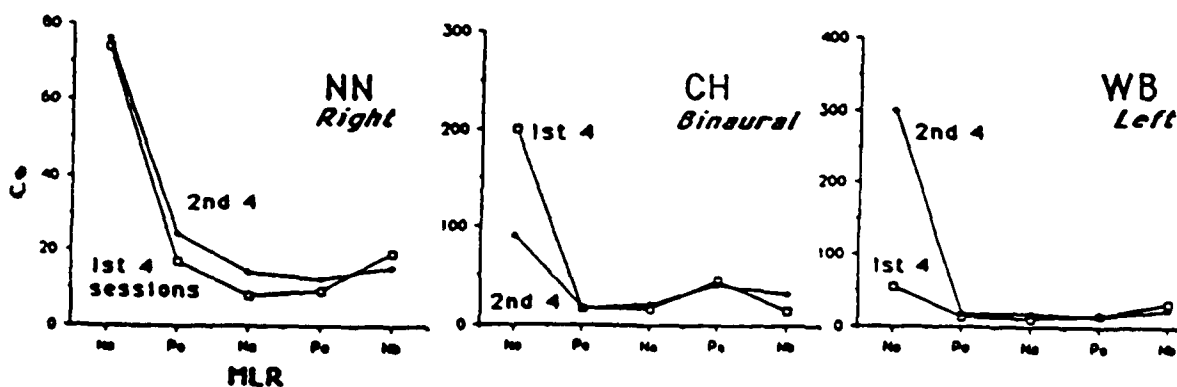


Figure 13. Panel A: Comparison of measures of absolute latency and latency stability for the middle latency response (MLR), parallel with Fig. 10 for the ABR. Panel B: Individual MLR stability profiles, indicating good replicability in profiles for first four weeks (open symbols) to second four weeks (filled symbols). Both panels illustrate the distinction between MLR peak No and later MLR peaks.

**A**

ABR and MLR latency stability
(8 Ss over 8 sessions)

Peak of EP

**B**

ABR, MLR, cortex latency stability
(ABR, cortex - 7 Ss;
MLR - 8 Ss)

Peak of EP
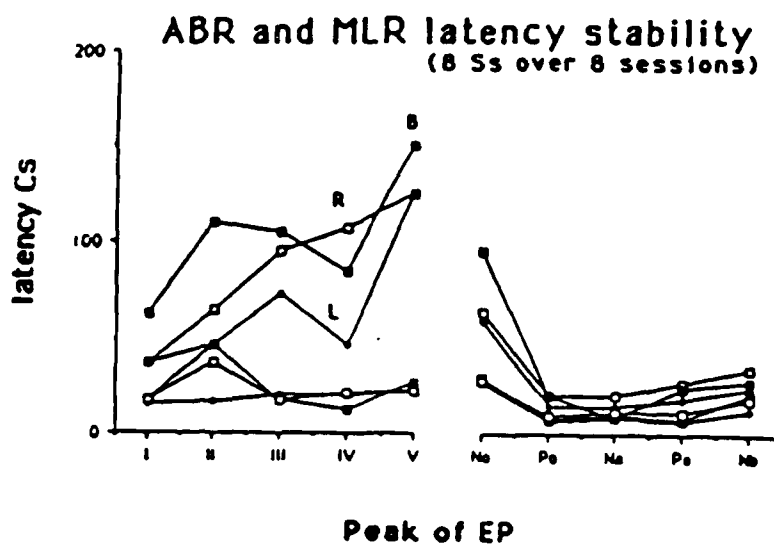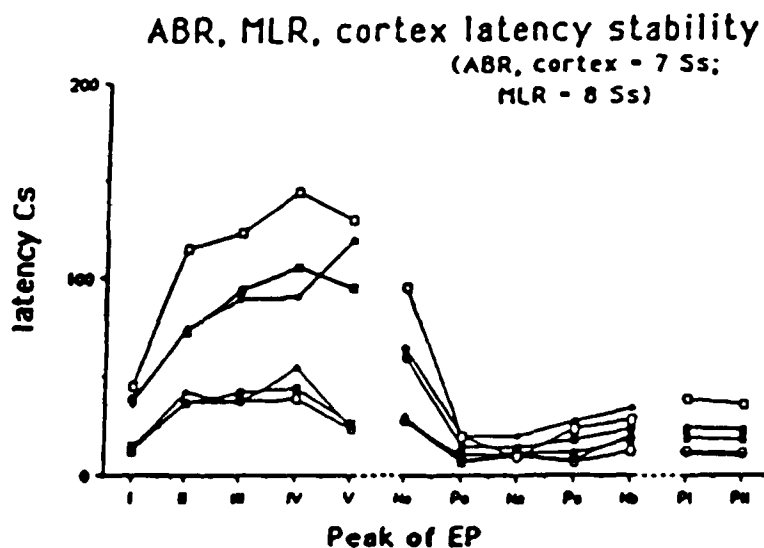
Figure 14. Panel A: Comparison of latency stability observed in groups of young adult subjects for ABR and MLR waveform peaks. Note the intermediate status of MLR peak No indicated by this measure. Panel B: Comparison of latency stability in young adults for ABR, MLR, and cortical auditory EPs. Note the similarity between ABR peaks and MLR peak No on the one hand, and the later MLR peaks and cortex on the other.
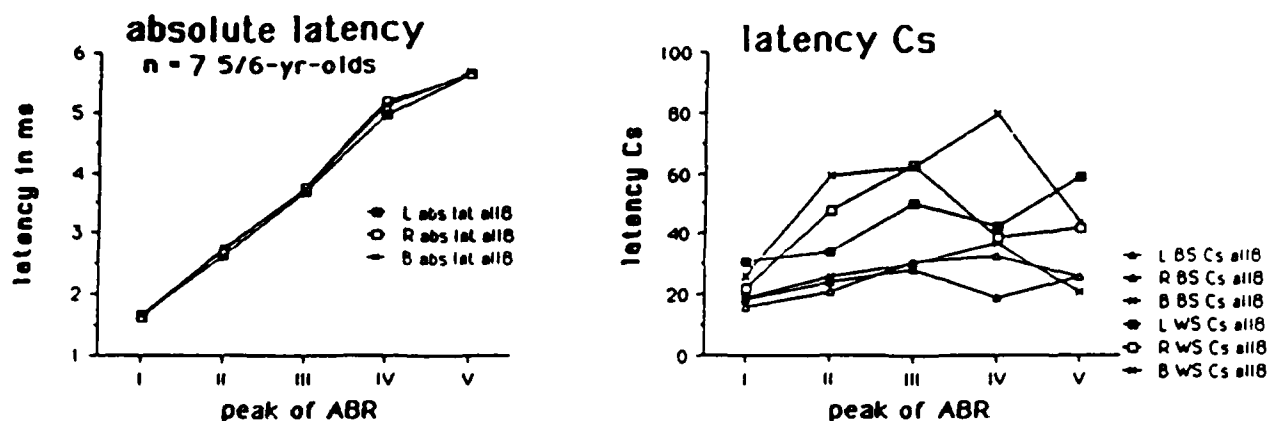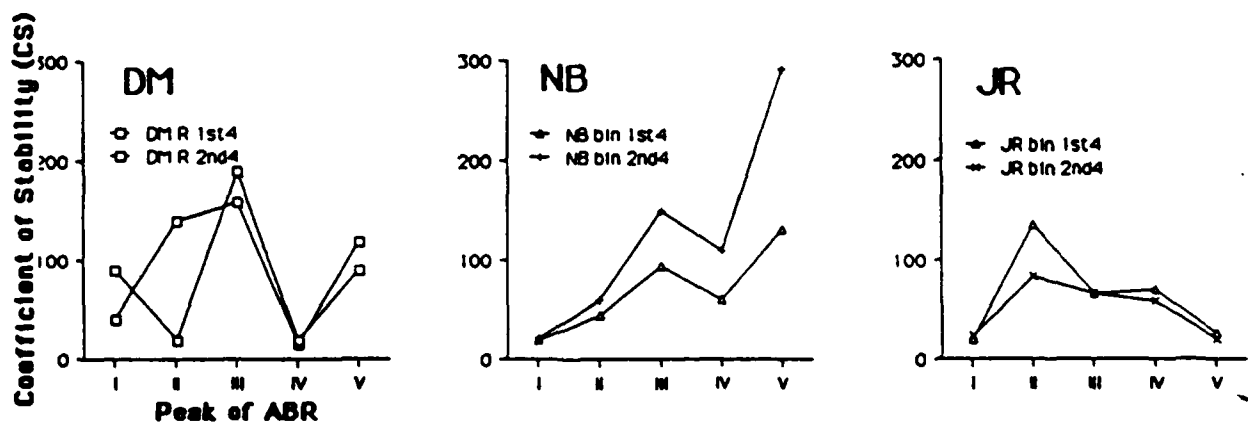
Figure 15. Panel A: ABR absolute latency and latency stability compared for a group of seven 5-6-year-old children. Panel B: ABR latency variability profiles for three individual 5-6-year-olds, showing good replicability from first four weeks to second four weeks of testing.
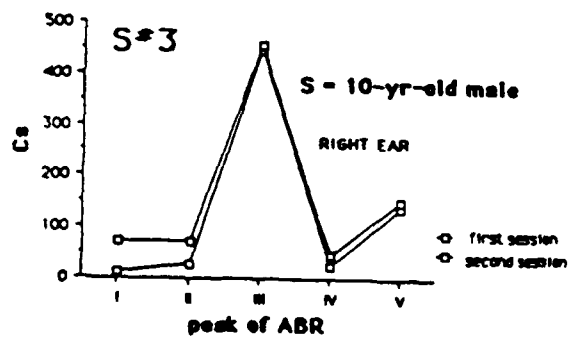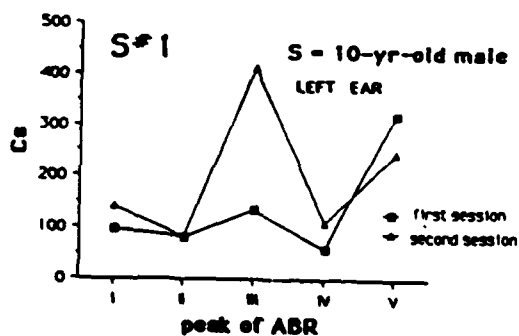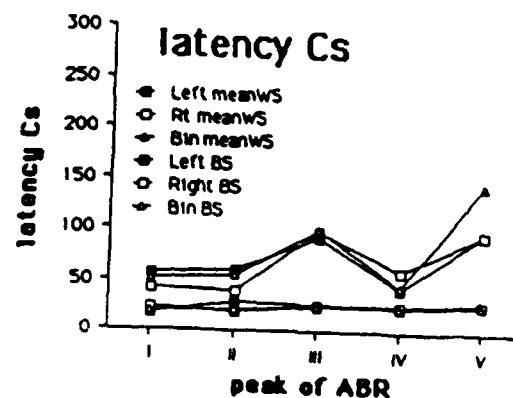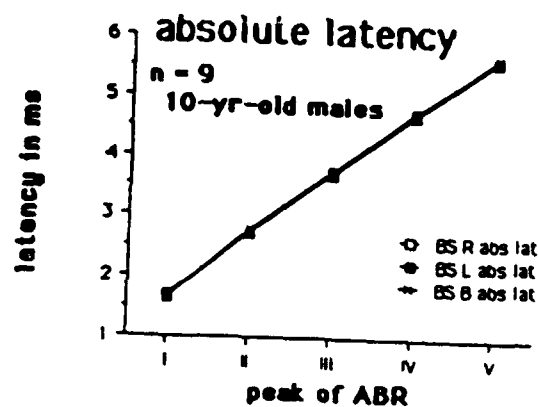
Figure 16. Panel A: ABR absolute latency and latency stability compared for a group of nine 10-11-year-old males. Panel B: ABR latency stability profiles for two individuals, showing good replicability from first four weeks to second four weeks of testing.

**A**

**B**
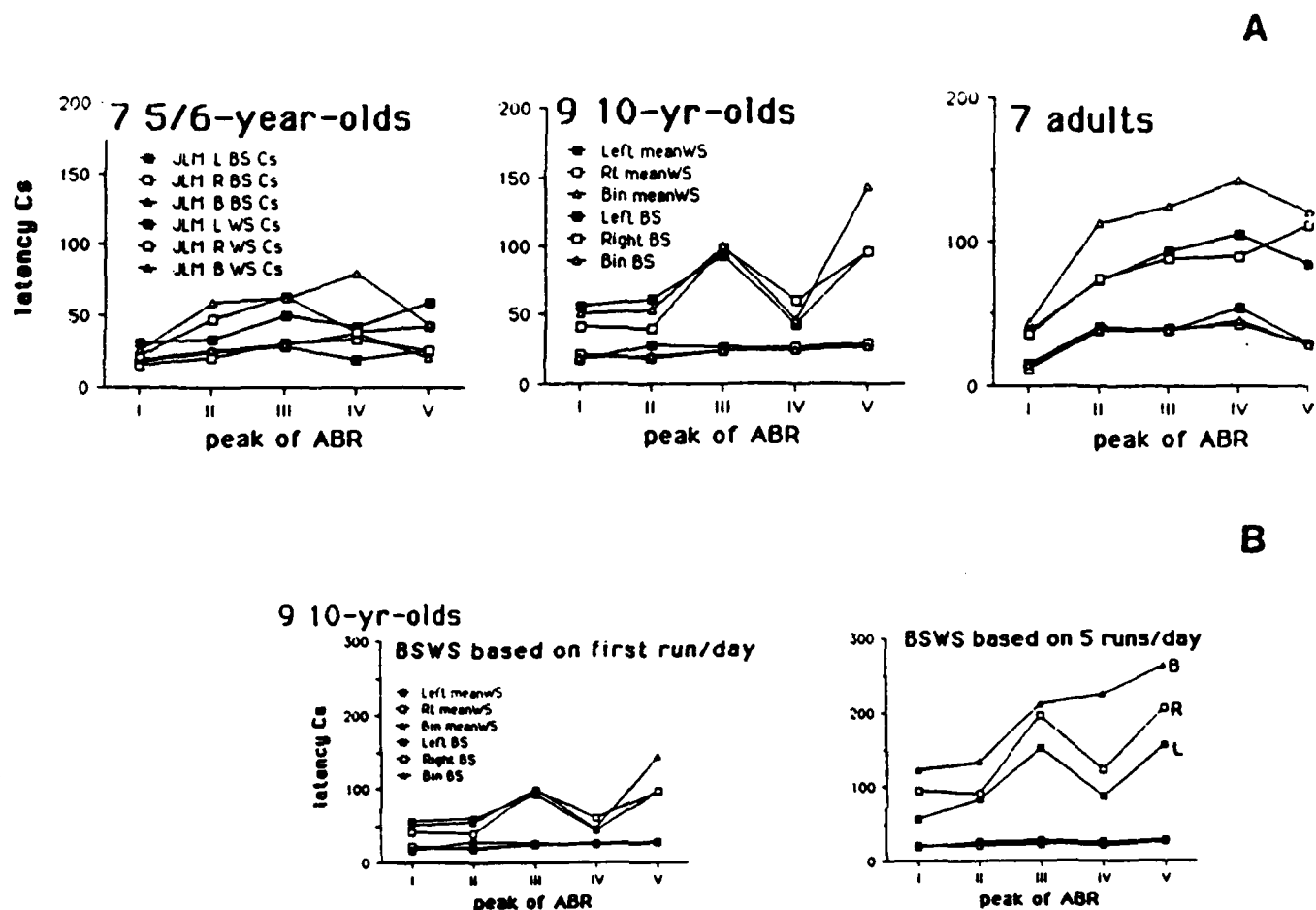
Figure 17. Panel A: Comparison of ABR latency stability for three age groups, showing a slight increase in between-subject consistency from children to adults, but a dramatic increase across age in the degree of within-subject consistency. Panel B: Differences in ABR latency stability as a result of time-base of comparison, with much higher consistency of responses seen for within-session repeated waveforms (on the right) than for between-session comparisons (on the left).
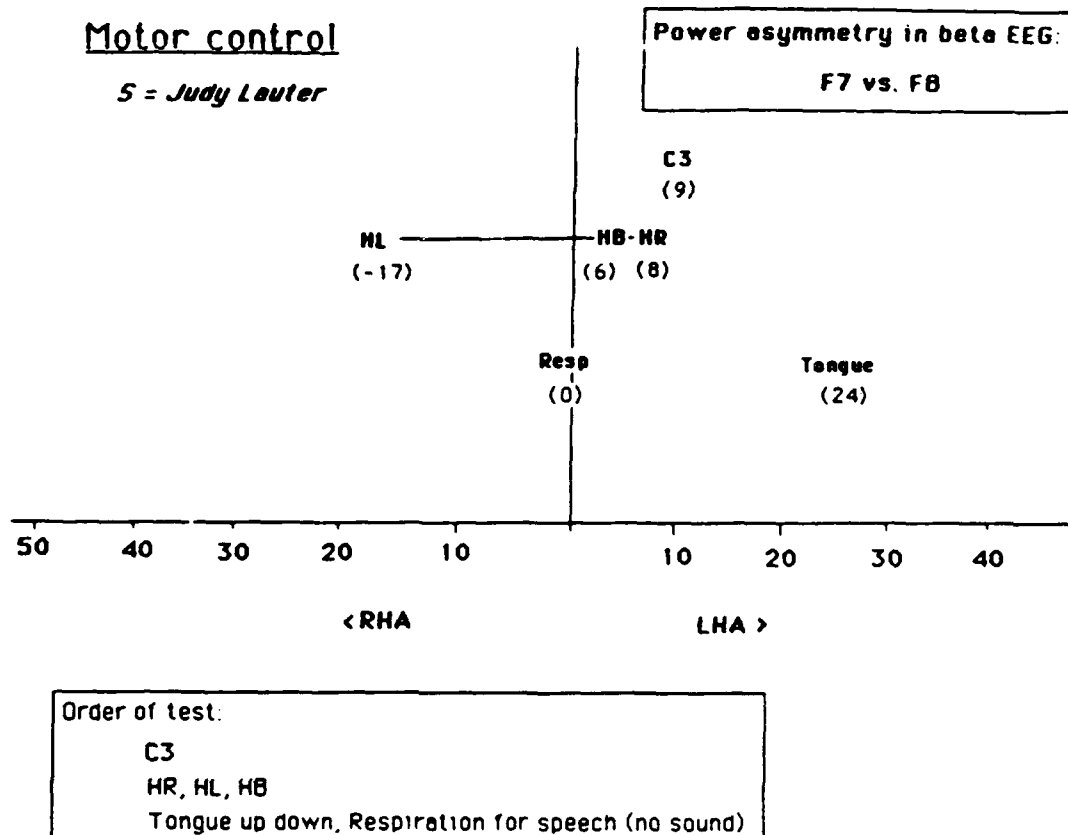
Figure 18. Values of EEG beta power asymmetry at electrode locations F7 and F8 calculated for each of six 5-min test conditions on one subject, plotted on a "relative hemisphere advantage" graph (analogous to Fig. 9): control ("C3"), unilateral right-hand movement ("HR"), unilateral left-hand movement ("HL"), bilateral hand movement ("HB"), tongue-tip movement ("Tongue"), and respiratory gestures during silent speech ("Resp"). Note: 1) the similarity between control and preferred-hand (right) and both-hand conditions, 2) the contralateral effect (Right Hemisphere Advantage RHA for left hand, LHA for right), and 3) the distinction in the two conditions related to speech motor gestures, with tongue-tip movement eliciting a large LHA and silent-speech respiratory movements showing no asymmetry at this electrode location.

# Complex sounds

*S = Judy Lauter*
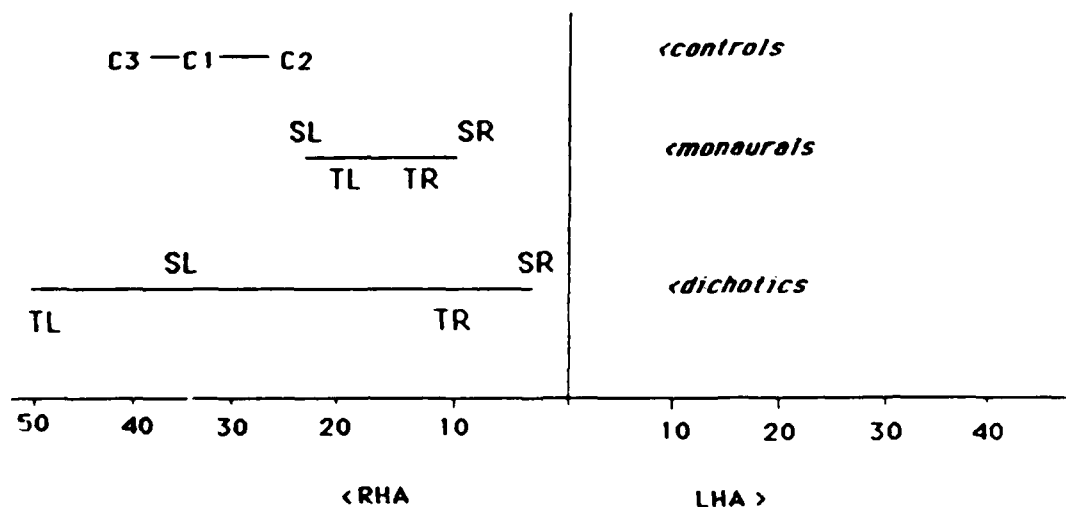
Power asymmetry in beta EEG.
T3 vs. T4

C3 —C1—— C2

         *‹controls*

SL_____SR

TL   TR

         *‹monaurals*

SL_____SR

TL              TR

         *‹dichotics*

50   40   30   20   10      10   20   30   40

‹ RHA             LHA ›

```
C = Control                          M = monaural
T = three-tone patterns              D = dichotic
S = synthetic syllables bdgptk +a
R = right ear (attended ear in dichotic)
L = left ear (‐)

                    order of test:
                      C 1
                      all syllables (m then d)
                      C 2
                      all tones (m then d)
                      C3
```

Figure 19. Values of EEG beta power asymmetry at electrode locations T3 and T4 calculated for each of eleven 5-min test conditions on one subject, plotted on a "relative hemisphere advantage" graph (cf. Figs. 9 and 18): three controls ("C1-3"), four monaural conditions (syllables "S" or tone patterns "T" in the right "R" or left "L" ears), and four dichotic conditions ("S" or "T" dichotically, with attention to "R" or "L" ears). Both monaural and dichotic conditions illustrate: 1) side-of-space asymmetry, with left-ear input eliciting a larger RHA than right-ear input, and v.v.; 2) asymmetry based on type of sound, such that--except for left-ear monaural--the tones always elicit a larger RHA than the syllables; and 3) interaction between these two, such that extreme asymmetries are seen for the combination of tone patterns with left-ear attention (largest RHA) vs. syllables with right-ear attention (smallest RHA).

<RHA                    LHA>*

15      10      5           5       10      15

T———S

T ——————————————— S                    JL

T - S

T—————— S                    CB

S——————————————— T    JM

S

S + T

DW

S

30      20      10          ;0      20      30      40

<LEA                    REA>
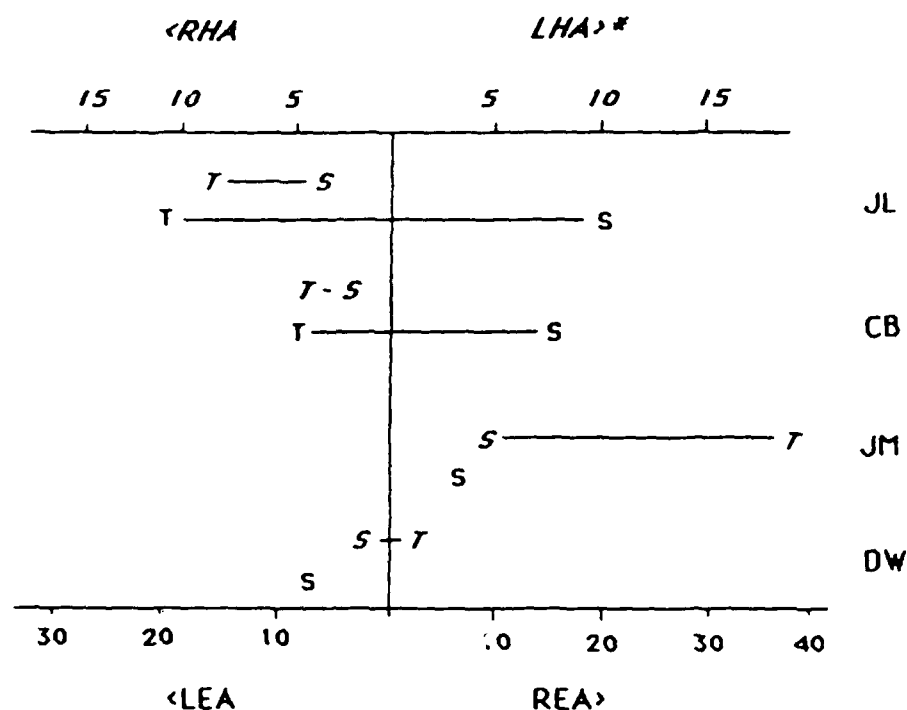
*calculated as shown in example below for JL

| conditions | JL T3 abs B | JL T4 abs B | JL hemis adv |
|---|---|---|---|
| 1 | SDR | 5.2 | 5.5 | 3 RHA |
| 2 | SDL | 5.0 | 10.3 | 5.3 RHA |
| 3 | | | | mid-2.8 RHA |
| 4 | TDL | 6.4 | 20.7 | 14.3 RHA |
| 5 | TDP | 9.4 | 11.8 | 2.4 RHA |
| | | | | mid-8.4RHA |

Figure 20. Comparison of asymmetries observed in four subjects with behavioral
(dichotic) testing (non-Italic symbols, lower abscissa) vs. qEEG testing (Italic symbols,
upper abscissa) on two sound sets, synthetic stop-CV syllables "S" and 200-ms IOI tone
patterns ("T"). Subjects JL and CB show good agreement between the relative
asymmetries in the two measures for the tone patterns and syllables, with tone patterns
eliciting left-ear-ward and right-hemisphere-ward dominance, and syllables the opposite.
Subjects JM and DW were not able to complete behavioral testing for the tone patterns, but
their data reveal good matches between the direction and magnitude of behavioral and
qEEG asymmetry elicited by the syllables.
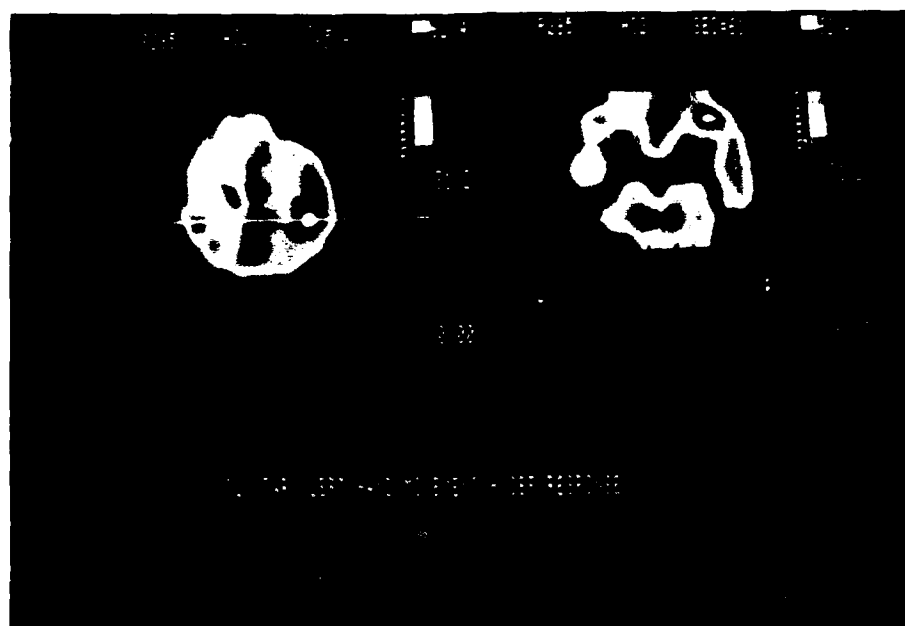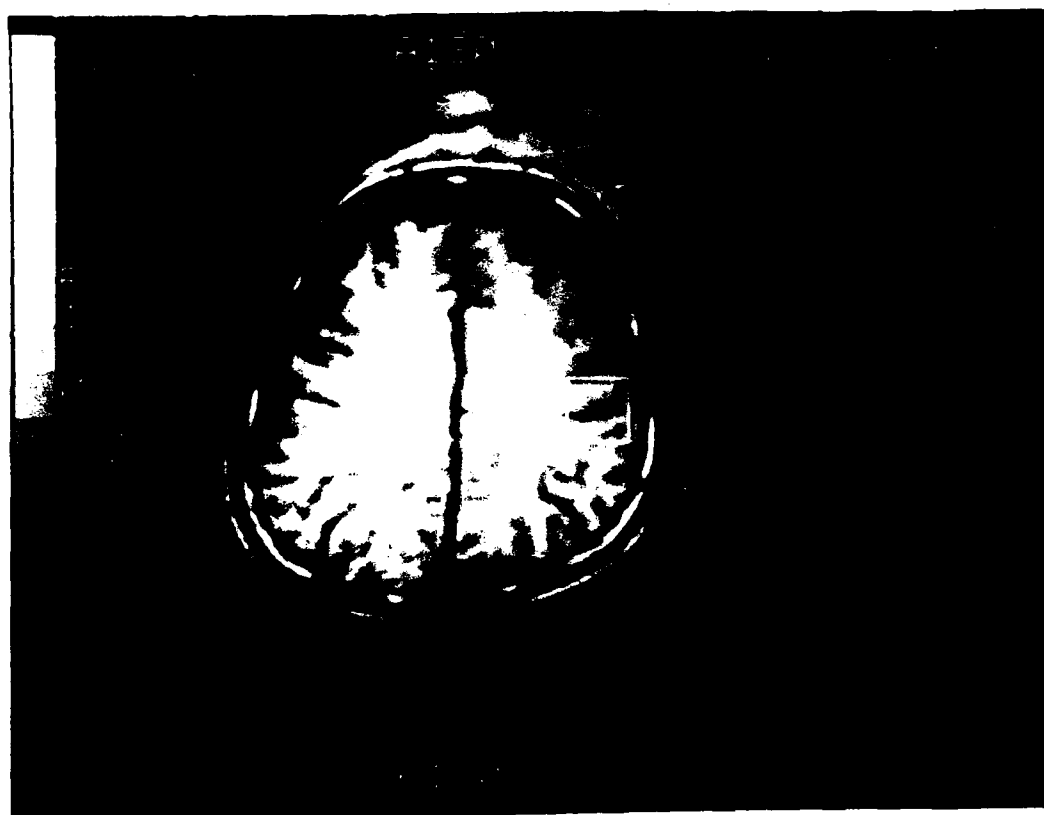
Figure 21. Panel A: Distribution of regional cerebral blood flow (rCBF) in the brain of a subject who is moving his left hand approximately 1/sec. The white region of activation appears in the cortical "hand area," whether viewed on a horizontal (left) or coronal (right) reconstruction. (Images generated on the Washington University PETT VI using i-v oxygen-15 and 40-sec scan time.) Panel B: White cursor outlining the region of interest (ROI) selected for our pilot experiment regarding the feasibility of using water-suppression volume-selection proton spectroscopy for studying physiological responses to hand movement. Placement of the cursor was guided in part by the PET results shown in Panel A, both for anterior-posterior location and rostral-caudal (in a slice just above the last appearance of the ventricles).

Figure 22. Proton spectra generated using water suppression, volume selection, and Fourier analysis with an MRI device, for five conditions tested on a single male subject in the following order: 1) baseline, 2) right-hand flexion, 3) control (no activation), 4) control (no activation), right-hand flexion. Peak at 4.7 PPM represents protons in water; values between +/-2 PPM represent protons in a variety of molecules, such as lactic acid, associated with neural activity.

Figure 23. Change spectra derived from tracings in Fig. 22: the baseline spectrum in Fig. 22 is subtracted from the spectrum from each of the test conditions to yield tracings showing differences comparing baseline and test: 1) right-hand flexion #1, 2) control #1, 3) control #2, 4) right-hand flexion #2. Panel A: positive changes relative to baseline; Panel B: negative changes relative to baseline.



Figure 24. Highlighted portions of change spectra from Fig. 23, showing only changes between +2 and -2 PPM, representing protons within molecules associated with neural activity. Panel A: positive changes relative to baseline; Panel B: negative changes relative to baseline.
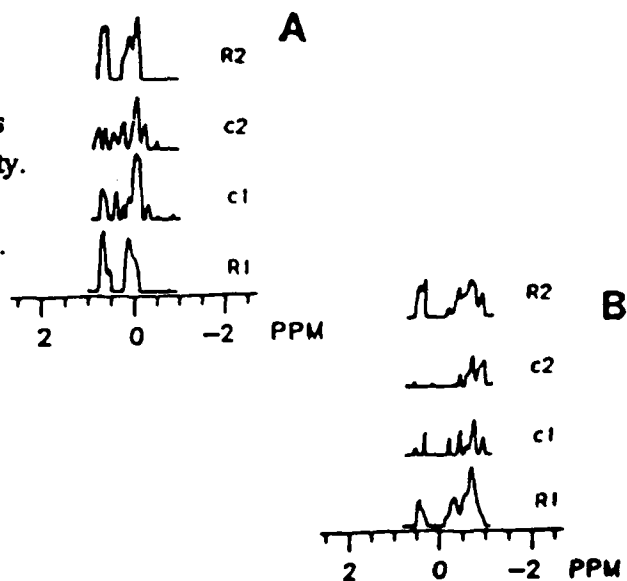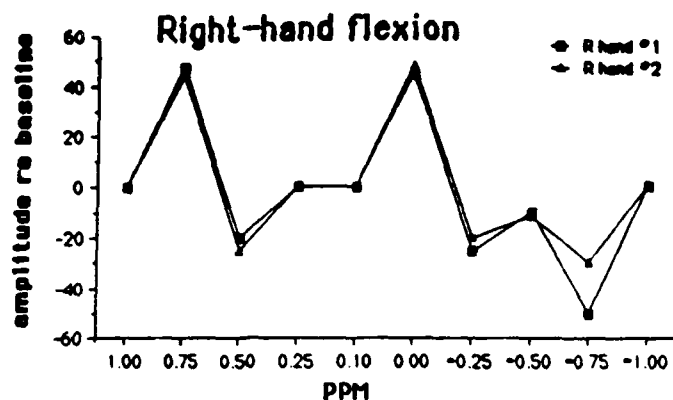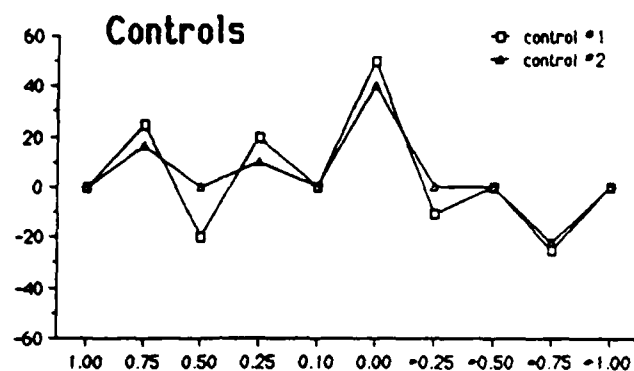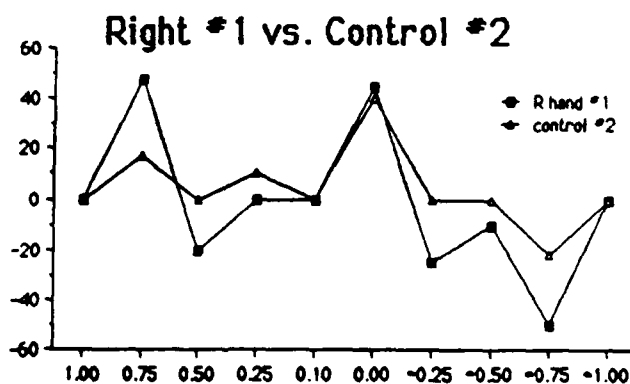
Figure 25. Schematics of difference spectra of Fig. 24, with amplitude of positive and negative changes graphed as a function of ten equally-spaced spectral locations. Panel A: changes from baseline for the two right-hand flexion conditions; Panel B: changes from baseline for the two control conditions (note how both positive and negative changes still apparent in the first control move closer to zero--i.e., baseline values--in the second control); Panel C: changes from baseline contrasting the two "extreme" conditions, the first right-hand flexion vs. the second control.
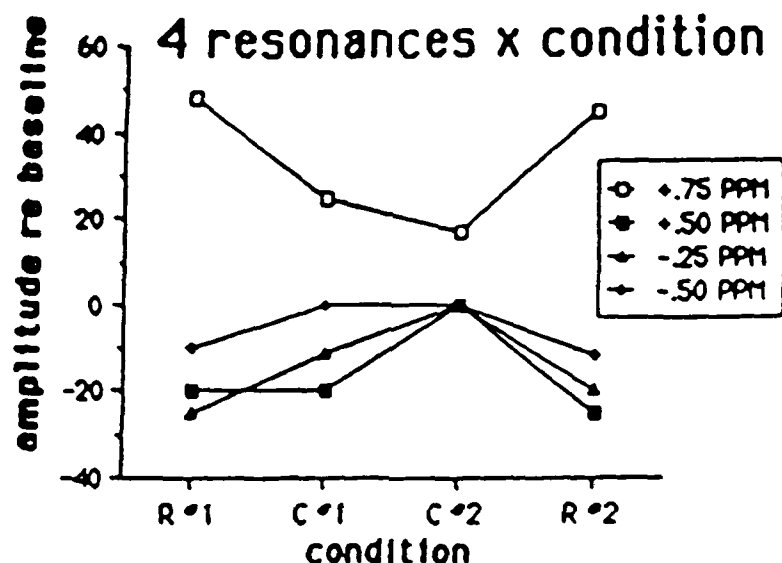
Figure 26. Amplitude of positive and negative spectral changes re baseline graphed as a function of test condition, with four selected spectral locations as the parameter. The graph illustrates how both positive and negative changes induced by the right-hand flexion drift toward zero through the two controls, to return to original activation levels with the second movement condition. Figures 22-26 suggest that: 1) hand movement results in both an increase in some chemicals and a depletion of others, 2) with time (sampled in the two controls), the chemicals which are increased with hand movement dissipate, and those which are depleted are restored, 3) for some molecules, these recovery processes are completed within the half-hour represented by the time required for the two control scans, while others may require a longer time to achieve baseline levels, and 4) a repetition of the movement condition, at least given the half-hour rest allowed here, can result in increases in the "positive-change" substances and depletion of the "negative-change" substances to the same degree as in the original movement condition.

A

B

$$T \text{ —C3—C1—C2———————S}$$

50   40   30   20   10        10   20

&lt; RHA                          LHA &gt;
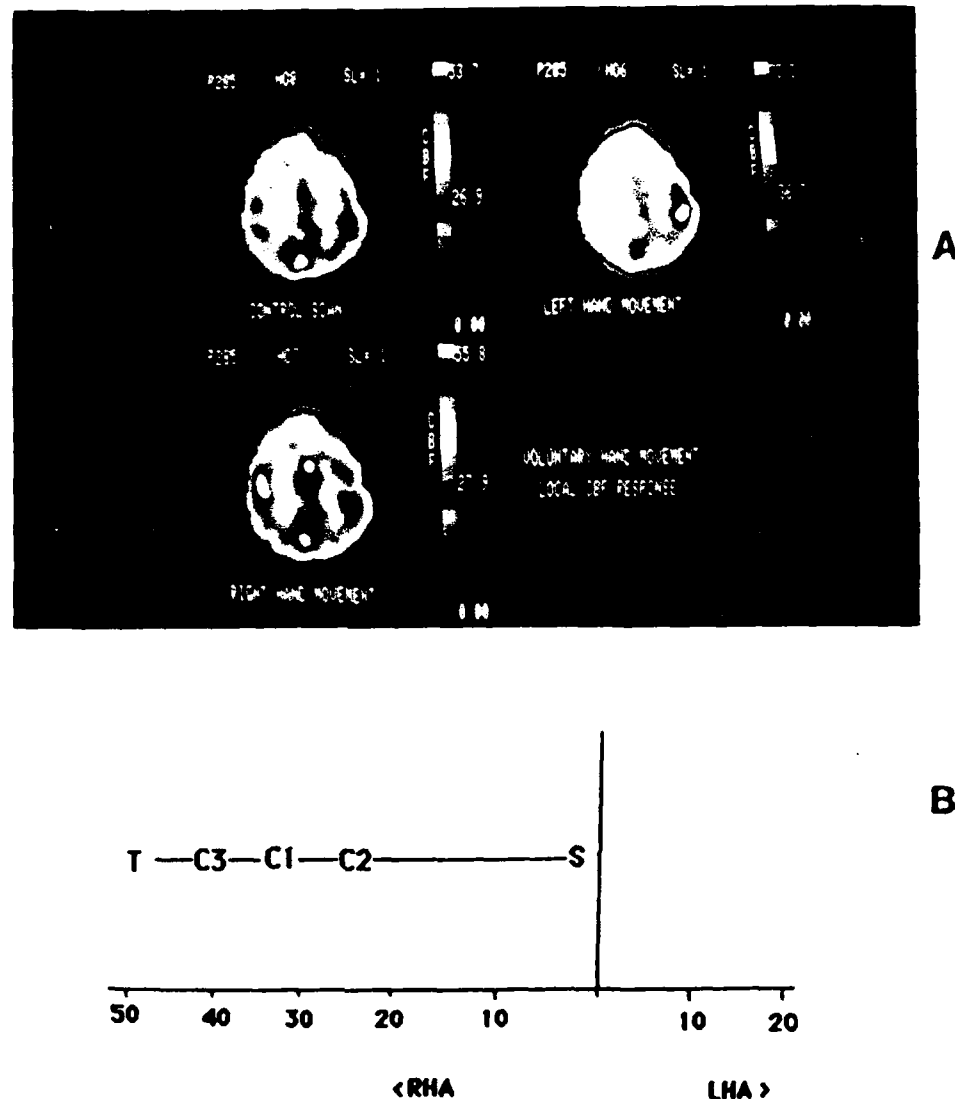
Figure 27. Evidence from other noninvasive methods of monitoring brain activity illustrating the same type of "residual activity" as observed in the MRS data of Figs. 22-26. Panel A: three PET images showing activation in response to hand movement, with chronology of the test conditions indicated by the degree of residual activity apparent on each image: left-hand first, right-hand second, and control third (control shows residual effects of both hand-movement conditions, more than one-half hour after the left-hand scan). (Images generated on the Washington University PETT VI using i-v oxygen-15 and 40-sec scan time.) Panel B: Selected data from a qEEG test with complex sounds (cf. Fig. 19), showing the beta power asymmetry at electrode locations T3/4 during two test and three control conditions, with chronology of the control conditions indicated by the shifts from baseline resting asymmetry ("C1") following testing with syllables ("S"--C2 moves toward LHA) and testing with tone patterns ("T"--C3 moves toward RHA).
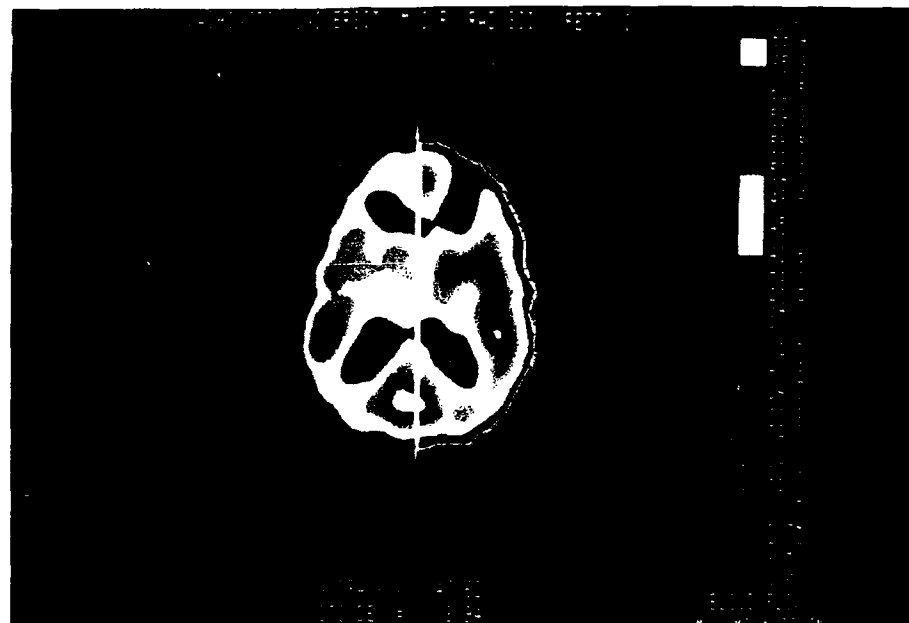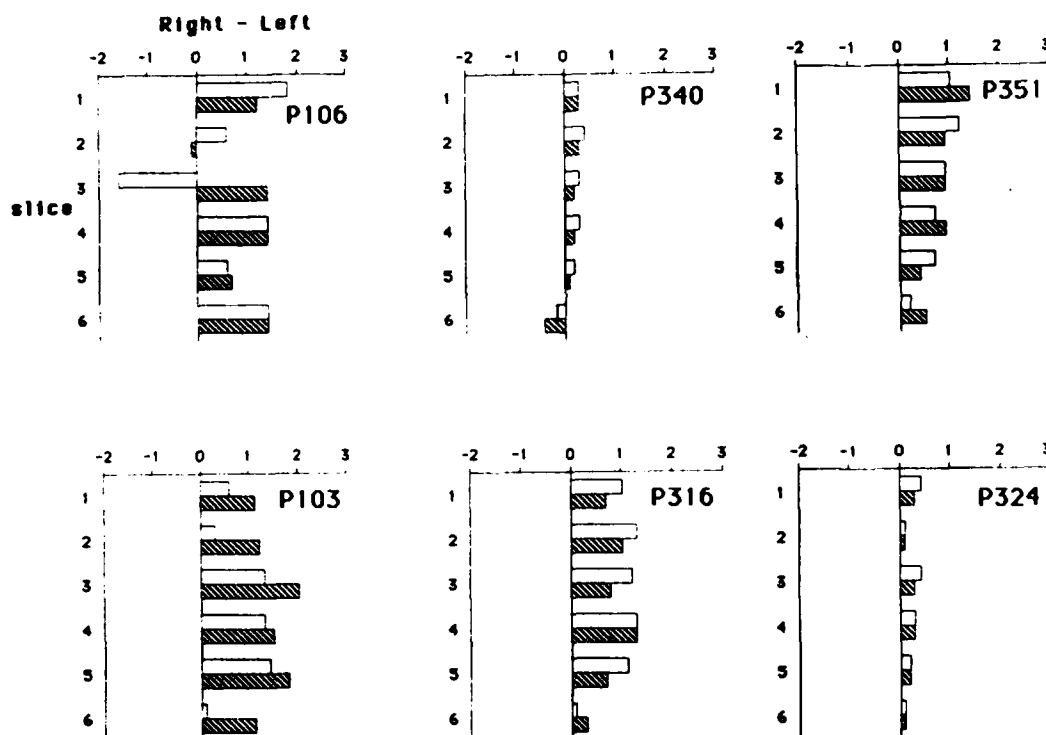
Figure 28. Quantification of resting regional cerebral blood flow (rCBF) in six subjects, comparing global right- vs. left-hemisphere rCBF. Panel A: illustration of measurement method, showing right hemisphere of slice 3 of subject P340 outlined in white, with mean and standard deviation rCBF within that outline displayed below the image. Panel B: comparison of global hemisphere rCBF values for each of six slices for each of six subjects, during first control scan (open bars) and second control scan (striped bars) from a single test session. Note the general predominance of right-hemisphere advantage in resting rCBF, and the good within-subject replicability of the degree of dominance from control scan #1 to control #2. Slices are numbered from 1 (most rostral) to 6 (most caudal). (Images taken with the Washington University PETT VI, using i-v oxygen-15 and 40-sec scan time.)

## APPENDIX

1. Lauter, J.L., Curriculum vita

2. Lauter, J.L., and Hirsh, I.J. (1985) Speech as temporal pattern: a psychoacoustical profile. Speech Communication 4: 41-54.

3. Lauter, J.L. (Ed.) (1985) Proceedings of the conference on the planning and production of speech in normal and hearing-impaired individuals: A seminar in honor of S. Richard Silverman. ASHA Reports 15.    [excerpts]

4. Lauter, J.L. (1985) Respiratory function in speech production by normally-hearing and hearing-impaired talkers: A review, In J. L. Lauter (Ed), [see #3], pp. 58-60.

5. Lauter, J.L., Herscovitch, P., Formby, C., and Raichle, M.E. (1985) Tonotopic organization in human auditory cortex revealed by positron emission tomography. Hearing Research 20: 199-205.

6. Lauter, J.L. and Loomis, R.L. (1986) Individual differences in auditory electric responses: Comparisons of between-subject and within-subject variability. I. Absolute latencies of brainstem vertex-positive peaks. Scandinavian Audiology 15: 167-172.

7. Lauter, J.L. and Pearl, N.B. (1986) Within-subject and between-subject measurements of stop-consonant productions by female talkers of English, Japanese, Navajo, and Spanish. Presented to Acoustical Society of America, Anaheim, November 1986. Abstract: Journal of the Acoustical Society of America 80: S62.

8. Lauter, J.L., and Karzon, R.G. (1987) Individual differences in auditory-evoked potentials: Variability of middle-latency responses, including comparisons with brainstem AEPs. Presented to Acoustical Society of America, Indianapolis, May 1987. Abstract: Journal of the Acoustical Society of America 81: S8.

9. Lauter, J.L., and Lu, F-L. (1987) VOT variability in Mandarin Chinese: Interactions with tone. Presented to International Society for Phonetic Sciences, Miami Beach, November 1987.

10. Lauter, J.L. (1988) Windows on the brain: what contemporary imaging devices can reveal about speech and hearing. Handbook for workshops given in Madison WI (September 1988) and Tucson AZ (February 1989)

11. Lauter, J.L. (1988) Functional-activation asymmetries in normal humans studied with quantitative EEG (qEEG): first tests in the CNS Project. Presented to Acoustical Society of America, Honolulu, November 1988. Abstract: Journal of the Acoustical Society of America 84: S57.

12. Lauter, J.L., and Loomis, R.L. (1988) Individual differences in auditory electric responses: Comparisons of between-subject and within-subject variability. II. Amplitude of brainstem vertex-positive peaks. Scandinavian Audiology 17: 87-92.

13. Lauter, J.L., Herscovitch, P., and Raichle, M.E. (1988) Human auditory physiology studied with positron emission tomography. In J. Syka and R.B. Masterton (Eds.), Auditory pathway (NY: Plenum Pub. Co.); pp. 313-317.

VITA

Judith L. Lauter, Ph.D.

## I. IDENTIFYING INFORMATION

Address: 7102 Avenida del Potrillo Tucson AZ 85747
Born: April 30, 1944 Austin, Texas
Social Security Number: 365-48-8763

Teaching Interests: Anatomy and physiology for speech & hearing,
Acoustic phonetics and speech perception, Sensory perception

Research Interests: Dichotic listening to speech and nonspeech
complex sounds, Coordinated noninvasive methods for studying
complex sensory function in the same subjects, including study
of indvidual differences and processing asymmetries (MRI, EPs,
qEEG, MEG, NMR-MRS, PET)

## II. EDUCATION

| | | |
|---|---|---|
| B.A. English | University of Michigan | 1966 |
| M.A. English | University of Arizona | 1968 |
| M.A. Information Science | University of Denver | 1971 |
| M.A. Linguistics | Washington University (St. Louis) | 1974 |
| Ph.D. Communication Sciences | Washington University (St. Louis) | 1979 |

## III. EXPERIENCE

1974-1979 Research Assistant, Central Institute for the Deaf

1979-1985 Research Associate, Central Institute for the Deaf,
and Assistant Professor of Communication Sciences,
Washington University at St. Louis

1985-1988    Associate Research Scientist, Dept. of Speech and
                      Hearing Sciences, University of Arizona

1988-          CNS Project Director, Institute for Neurogenic Communication
                      Disorders, University of Arizona

## IV. PROFESSIONAL ACTIVITIES

Professional Associations:

Acoustical Society of America
International Society of Phonetic Sciences
Society of Sigma Xi

Consultantships:

McDonnell Center for the Study of Higher Brain Functions,
Washington University at St. Louis (1981-1985)

Journal of Speech and Hearing Research, ad hoc reviewer

Journal of the Acoustical Society of America, ad hoc reviewer

Psychobiology, ad hoc reviewer

NINCDS site visit for U. Minnesota PPG proposal on Communication
Disorders, 1983

Collaborator status with Los Alamos National Laboratory:
Life Sciences Division, Neural and Biological Sciences
Group (initiated April 1987)

PET Steering Committee for University of Arizona

<u>Instructional Contributions</u>:

Responsible for the following courses at Washington University
in St. Louis:

SpH 401   Anatomy and Physiology for Speech and Hearing
SpH 433   Acoustic Phonetics and Speech Perception
  [team-taught with I.J.Hirsh]

Diagnostic Audiology: lectures on auditory physiology
        and central auditory disorders
Psychology of Speech and Language: lectures on aphasia
Psychology of Music: lectures on temporal perception
Basic Sciences course for otolaryngology residents:
        lectures on auditory anatomy and physiology
Neurological Pathophysiology course for second-semester
Washington University School of Medicine
        students: lectures on auditory function
        and disorders


Course contributions at University of Arizona:

Neurological Foundations of Psychology (Psych 302):
    lecture on PET as a window on normal human physiology
Introduction to Biopsychology (Psych 403):
    lectures on development and noninvasive study of
    speech & hearing system in humans
Research and Theory in Biopsychology (Psych 520); lecture
    on modern imaging techniques, including PET-scan
Speech Science (S&H 260): lecture series on methods for
    studying neurological foundations of speech perception

## V. RESEARCH ACTIVITIES

Directed studies and thesis committees:

(Washington University, St. Louis):

Discrimination of Multisyllabic Sequences by Young
Infants (Roanne Karzon, Ph.D. Dissertation, 1982)

Hypnosis and Its Relationship to Right Hemisphere
Processing (Jeffrey Levine, Ph.D. Dissertation, 1983)

Perceptual Attributes of Babbling From Four to Twelve Months
(Ginger Kuehn, Ph.D. Dissertation, 1984)

Dimensional Tradeoffs in the Perception of Complex
Tone Sequences (Punita Singh, M.A. Thesis, 1984)

(University of Arizona, Tucson)
[All students are S&H unless otherwise noted]

Independent Studies:
Carol Baldwin [Psychology], "Readings in acoustic phonetics"
(completed)
Nancy Pearl, "VOT variability: a cross-language study"
(completed; paper given at ASA in Anaheim fall 1986;
MS submitted)
Fang Ling Lu, "VOT production in Chinese talkers:
comparisons between Chinese and English stops"
(completed; paper given at ASA in Miami fall 1987; MS in
prep)
Julie Petersen, "Readings in auditory physiology as
background to studies of auditory CNS asymmetries with
PET" (in progress)

Master's Projects:

> Ron Mack (R. Curlee, Advisor) "Reaction time measures in
> dysfluents and normals" (completed)

Ph.D. Internships:

> Elena Plante (L. Swisher, Advisor) "Regional cerebral blood
> flow activation asymmetries in human brains during
> rest and auditory stimulation" (in progress)

Ph.D. Dissertations:

> Carol Baldwin [Psychology; M. Wetzel, Advisor], "The voice
> of emotion: acoustic properties of six emotional
> expressions" (completed)
> Janet Lord [Educational Psychology; S. Mishra, Advisor],
> "Short latency evoked potentials and intra-individual
> variability in children" (completed)
> Robert Oyler [N. Matkin, Advisor], "Individual differences
> in ABRs in ten-year-old males" (in progress)

Sponsored Projects:

### Funded

| | | | | |
|---|---|---|---|---|
| AFOSR 84-0335 | 84-85 | "Complex sounds" | 100% | $71,000 |
| AFOSR 85-0379 | 85-88 | "Dichotic listening..." | 100% | 307,000 |
| AFOSR 87-0003 | 86 | "PET satellite station" | -- | 80,500 |
| AFOSR 88-0352 | 88-91 | "CNS Project" | 100% | 450,000 |
| ASHF/Apple | | "CAD tests" (equipment) | -- | 15,000 |

### Approved, not funded

| | | |
|---|---|---|
| NIH (NINCDS) | "Individ. diffs in EPs" | 650,000 |
| NIH (NINCDS) | "PET satellite system" | 380,000 |

## VI. LECTURES AND PAPERS  (Invited and Submitted)

### 1979

"A speech microscope," (with R Vemula, AM Engebretson, R Monsen) to
Acoustical Society of America, Cambridge.

### 1980

"Dichotic identification of complex sounds," to Acoustical
Society of America, Atlanta. Abstract: J Acoust Soc Amer 67:
S100.

### 1981

"Dichotic listening reconsidered as a type of masking paradigm," to
Acoustical Society of America, Ottawa. Abstract: J Acoust Soc
Amer 69: S22.

### 1982

"Contralateral interference and relative ear advantages for event
timing in three-tone patterns," to Acoustical Society of America,
Chicago. Abstract: J Acoust Soc Amer 71: S47.

"The psychophysics and neurophysiology of hearing," lecture series
presented to the Brain Breakfast Club, Washington University
Department of Neurology, St. Louis.

"Ear advantages: what are they good for?" to staff of Kresge Hearing
Institute of the South, New Orleans.

### 1983

"Bandwidth of three-element patterns and its effect on relative
ear advantages," to Acoustical Society of America, Cincinnati.
Abstract: J Acoust Soc Amer 73: S43.

"Human auditory cortex: a preliminary report on studies using
positron emission tomography (PET)," (with P Herscovitch and
ME Raichle) to Acoustical Society of America, Cincinnati.
Abstract: J Acoust Soc Amer 73: S60.

"Cerebral metabolic effects of auditory stimulation," to Brain
Breakfast Club, Washington University Department of
Neurology, St. Louis.

"Tonotopic organization in human auditory cortex as revealed by
regional changes in cerebral blood flow," (with C Formby, P Fox,
P Herscovitch, ME Raichle) to XI International
Symposium on Cerebral Blood Flow and Metabolism, Paris.
Abstract: J Cereb Blood Flow Metab 3: S248-249.

"Dichotic listening and models of the nervous system," to Computer
Systems Group, Los Alamos National Laboratory, Los Alamos NM.

"PET and the cortex: the effects of auditory stimulation on
cerebral blood flow," to Department of Speech and Hearing
Sciences, University of Arizona, Tucson.

1984

"Changes in human regional cerebral blood flow in response to
pure tones," to Acoustical Society of America, Norfolk. Abstract:
J Acoust Soc Amer 75: S14.

1985

"Individual differences in the perception of frequency changes
in three-element sequences," to Acoustical Society of America,
Austin. Abstract: J Acoust Soc Amer 77: S36.

"Individual differences in auditory evoked responses: comparisons
of between-subject and within-subject variability in brainstem
and cortical waveforms," to Acoustical Society of America,
Austin. Abstract: J Acoust Soc Amer 77: S65.

"Human auditory central nervous system: visualization of cortical
and subcortical centers with regional cerebral blood flow
measurements on the PETT-VI," to Acoustical Society of America,
Austin. Abstract: J Acoust Soc Amer 77: S94.

"Workshop on Central Auditory Processing: Basic Science Background and
Clinical Realities", sponsored by the Speech and Hearing
Association of Greater St. Louis.

1986

"Dichotic listening: the good news and the bad news," Speech and
Hearing Sciences Colloquium, University of Arizona

"Individual differences in evoked responses: second report," to
Acoustical Society of America, Cleveland. Abstract: J Acoust
Soc Amer 79: S5.

"VOT variability: Within-subject and between-subject measurements of
stop-consonant production by female talkers of English, Japanese,
Navajo, and Spanish" (with N. Pearl), to Acoustical Society of
America, Anaheim. Abstract: J Acoust Soc Amer 80: S62.

1987

"Using the PET-scan to study normal human brain function," to Los
Alamos National Laboratory, Life Sciences Division, Los
Alamos NM, March 1987.

"Individual differences in auditory evoked potentials: middle-latency
responses," (with R. Karzon) to Acoustical Society of America,
Indianapolis. Abstract: J Acoust Soc Amer 81: S8.

"Human auditory physiology studied with positron emission tomography,"
(with P. Herscovitch and M.E. Raichle) invited paper to Auditory
Pathway: Structure and Function conference, Prague.

"VOT variability in stop-consonant productions by bilingual speakers of English and Mandarin Chinese," (with Fang-Ling Lu) to Acoustical Society of America, Miami. Abstract: J Acoust Soc Amer 82: S115.

"VOT variability in Mandarin Chinese: interactions with tone," (with Fang-Ling Lu) to International Society for Phonetic Sciences, Miami Beach.


1988

"The PET scan as a tool for studying human neurophysiology," to University of Arizona Medical School Neurology Grand Rounds, Tucson AZ.

"Gender differences in the production of vocal emotional expressions" (with Carol M. Baldwin and Peter C. Facciola), to Western Psychological Association, San Francisco CA.

"Windows to the brain," to American Speech-Language-Hearing Foundation Technology Conference, Mesa AZ.

"Toward a taxonomy of vocal expressions of emotion" (with Carol M. Baldwin and Peter C. Facciola), to Society for Psychotherapy Research, Santa Fe NM.

"Windows to the brain: What contemporary imaging devices can reveal about speech and hearing," University of Wisconsin University Extension Program Communication Disorders Workshop, Madison WI.

"Assessment techniques and what they can tell us," to American Speech-Language-Hearing Association Clinical Neuroscience Conference, Bethesda MD.

"Positron emission tomography as a tool for studying normal human brain function," to University of Tennessee PET Laboratory, Knoxville TN.

"Functional-activation asymmetries in normal humans studied with
    quantitative EEG (qEEG): first tests in the CNS Project,"
    to Acoustical Society of America, Honolulu HI. Abstract:
    J Acoust Soc Amer 84: S57.

1989

"Introduction to the new noninvasive techniques, from EPs to PET,"
    to Dept. of Psychiatry, University of Arizona Medical Center,
    Tucson AZ (Jan)

"Windows on the brain: What the new imaging techniques can tell us
    about speech, language, and hearing," Workshop sponsored
    by Dept. of Speech & Hearing Sciences, Tucson AZ (Feb)

"Comparisons of between- and within-subject variability in repeated-
    measures auditory brainstem responses (ABRs) in 5-6-year-old
    children" (with J.M. Lord-Maes), to Acoustical Society of
    America, Syracuse NY (May)

"Global brain asymmetries in regional cerebral blood flow (rCBF)
    during resting conditions with positron emission tomography
    (PET): establishing a baseline for experiments on brain
    asymmetries and complex sounds in the CNS Project" (with
    E. Plante), to Acoustical Society of America, Syracuse NY (May)

## VII. PUBLICATIONS

### 1979

Vemula, R., Engebretson, A.M., Monsen, R., and Lauter, J.L. A speech microscope. In J.L. Wolf and D.H. Klatt (Eds), Speech communication papers presented at the 97th meeting of the Acoustical Society of America. NY: Acoust Soc Amer; pp. 71-74.

### 1981

Lauter, J.L. Book review of JM Pickett, The sounds of speech communication [1980]. Ann Otol 90: 302.

### 1982

Lauter, J.L. Dichotic identification of complex sounds: absolute and relative ear advantages. J Acoust Soc Amer 71: 701-707.

### 1983

Lauter, J.L. Stimulus characteristics and relative ear advantages: a new look at old data, J Acoust Soc Amer 74: 1-17.

Lauter, J.L. Book review of NJ Lass (Ed.), Speech and language. Advances in basic research and practice [1981]. Volta Review 85: 42-43.

Lauter, J.L. Book review of NJ Lass et al (Eds.), Speech, language, and hearing, vols. 1-3 [1982]. New Engl J Med 307: 900-901.

### 1984

Lauter, J.L. Auditory system. In A.L. Pearlman and R.C. Collins (Eds.), Neurological pathophysiology, 3d. ed., London: Oxford Univ. Press; pp. 86-109.

Lauter, J.L. Contralateral interference and ear advantages for identification of three-element patterns. Brain and Cognition 3: 259-280.

Levine, J.L., Kurtz, R.L., and Lauter, J.L. Hypnosis and its effect
on left and right hemisphere activity. Biol Psychiat 19:
1461-1475.

1985

Lauter, J.L., and Hirsh, I.J. Speech as temporal pattern: a
psychoacoustical profile. Speech Communication 4: 41-54.

Lauter, J.L. (Editor) The planning and production of speech: report of
the conference. ASHA Report #15.

Lauter, J.L. Respiratory function in speech production by
normally-hearing and hearing-impaired talkers: a review. In J.L.
Lauter (Ed.), The planning and production of speech: report of
the conference. ASHA Report #15.

Lauter, J.L., Herscovitch, P., Formby, C., Raichle, M.E. Tonotopic
organization in human auditory cortex revealed by positron
emission tomography. Hearing Research 20: 199-205.

Lauter, J.L. Book review of FH Duffy and N Geschwind (Eds.), Dyslexia:
a neuroscientific approach to clinical evaluation [1985], and BP
Rourke (Ed.), Neuropsychology of learning disabilities [1985].
New Engl J Med 313: 898.

1986

Lauter, J.L., and Loomis, R.L. Individual differences in auditory
electric responses: comparisons of between-subject and
within-subject variability. I. Absolute latencies of brainstem
vertex-positive peaks. Scandinavian Audiology 15: 167-172.

Lauter, J.L. Book review of I. Reinvang, Aphasia and brain
organization [1985]. New Engl J Med 315: 268.

1988

Lauter, J.L. and R.L. Loomis. Individual differences in auditory
electric responses: comparisons of between-subject and
within-subject variability. II. Amplitudes of brainstem
vertex-positive peaks. Scandinavian Audiology 17: 87-92.

Lauter, J.L., P. Herscovitch, and M.E. Raichle. Human auditory
physiology studied with positron emission tomography. In
J. Syka and RB Masterton (Eds.), Auditory Pathway. Plenum:
NY; pp. 313-317.

Lauter, J.L. Book review of D. Bowsher, Introduction to the anatomy
and physiology of the nervous system [5th ed., Blackwell: Oxford,
1988]; and R. N. Rosenberg and A. E. Harding (Eds), The molecular
biology of neurological disease [Butterworths: London, 1988].
New Engl J Med 319: 875-876.

In press

Lauter, J.L. Auditory system. In A. Pearlman and R.C. Collins
(Eds.), The Neurobiology of Disease. Oxford Univ. Press.

# SPEECH AS TEMPORAL PATTERN: A PSYCHOACOUSTICAL PROFILE*

Judith L. LAUTER and Ira J. HIRSH

*Central Institute for the Deaf and Washington University, St. Louis, Missouri 63110, USA*

**Abstract.** In visual representations of the acoustical signal of speech, such as the waveform or spectrogram, speech appears as a series of concatenated sequences of acoustical events, which vary in spectrum, amplitude and duration. The results of a variety of psychoacoustical experiments, from auditory fusion to temporal masking to studies of streaming, can be interpreted as relevant to discovering the auditory capabilities used in listening to these speech sequences. A sampling of such results serves to illustrate the connections between the psychoacoustics of speech and nonspeech, and to suggest guidelines for future work on non-speech temporal patterns, with the goal of a more complete psychophysics of complex sounds.

**Zusammenfassung.** Im visuellen Abbild des akustischen Sprachsignals, etwa in den Kurven des Zeitsignals oder im Spektrogramm, tritt uns die Sprache als Folge untereinander verbundener akustischer Erscheinungen entgegen, die bezüglich ihres Spektrums, ihrer Amplitude und ihrer Dauer variieren. Aus den Ergebnissen einer Reihe psychoakustischer Experimente (from auditory fusion to temporal masking to studies of streaming) lassen sich wichtige Hinweise für die Einsicht in die auditorischen Fähigkeiten bei der Aufnahme solcher Sprachsequenzen ableiten. Die Auswertung dieser Ergebnisse kann zur Illustration der Verbindungen zwischen der Psychoakustik sprachlicher und nicht-sprachlicher Signale dienen und Richtlinien für die zukünftige Arbeit an der Zeitstruktur nicht-sprachlicher Signale erzeugen mit dem Ziel einer umfassenderen Einsicht in die Psychophysik komplexer Klänge.

**Résumé.** Dans les représentations visuelles du signal acoustique de la parole telles que l'oscillogramme ou le spectrogramme, la parole apparaît comme des séries concaténées de phénomènes acoustiques qui varient dans leur spectre, en amplitude et en durée. Les résultats d'une grande variété d'expériences psychiacoustiques, de la fusion auditive au masquage temporel en passant par les études sur la fluence (streaming) peuvent être interprétées comme pertinentes pour la découverte des capacités auditives utilisées dans l'écoute de ces séquences de parole. Un échantillonnage de ces résultats expérimentaux servira à illustrer les relations entre la psychoacoustique des phénomènes non verbaux et ceux de la parole de manière à suggérer des lignes de conduite pour le travail futur sur les configurations temporelles non verbales afin d'aboutir à une psychophysique plus complète des sons complexes.

**Keywords.** Auditory fusion, temporal masking, streaming, psychoacoustics of speech, non-speech temporal patterns, psychophysics of complex sounds.

## Introduction

In looking at representations of the acoustical signal of speech, one notices a number of characteristics which can be measured in terms of those acoustical dimensions, like intensity, frequency, and duration, which have been traditionally important in the study of hearing. The obvious question arises as to how these different characteristics are perceived by observers, not only when the dimensions describe sounds identifiable as speech, but also when similar characteristics occur in sounds that a listener would not associate with known speech forms. Although speech perception is generally understood as the science of how listeners perceive sounds-heard-as-speech, we believe that there is much to be discovered about auditory-system function for speech perception

that can be studied most effectively with sounds that listeners do not hear as speech-like. In order to illustrate the grounds and applications of this belief, we present here a *topical sampling of* psychoacoustical experiments using nonspeech stimuli to show how results from such experiments bear on the signal-processing problems that speech presents to a listener.

## The temporal structure of speech

### Speech segments in waveform and spectrogram

One of the most obvious characteristics of visual representations of the speech signal, from the gramophone tracings published by E.W. Scripture in 1902 to modern spectrograms and oscillograms, is its segmental quality. L.A. Chistovich (e.g., [9]) has long been concerned with the relation between speech perception and the "segmentation cues" that seem to be distributed throughout the speech signal. Both waveforms and spectrographic representations show that the acoustical "speech stream" is not a smooth-flowing, homogeneous ribbon of sound, but rather comprises sequences of intermittent sound elements that differ from one another along a number of acoustical dimensions.

For example, the waveform reveals a succession of segments that differ from each other in amplitude, spectrum, and duration. The alternation of higher and lower amplitude portions of the signal can be shown to correlate with changes from vowels (higher-amplitude regions) to consonants (lower-amplitude portions of the wave). This amplitude modulation (AM) ranges from 100% (the *silence preceding the burst of plosives*) to lesser amounts, and the range of sound intensities covers a range of 40 dB for inter-phonemic (or "segmental") distinctions, e.g., between the peak of /a/ and the noise of /f/ or a /p/ burst. The dynamic range of sequential amplitude changes is even larger if we consider narrative speech, where prosodic shaping of the speech signal ("supraseg- mental" changes) can cause some vowel peaks to rise even further above the phonetic regions of silence.

Spectral changes from point to point in the speech signal also can be clearly seen in the waveform. Fine timbral distinctions, such as the spectral differences between vowels, are hard to see, but the points of change between periodic and aperiodic sounds can be easily defined and measured. A number of durational aspects can be monitored with the waveform display, as well: from timing of the brief noise bursts of stop and affricate consonants, to the rhythmic units of duration measured from peak to peak of vowel amplitude.

The spectrograph's design as a set of filters dictates that its display will focus on the spectral characteristics of speech segments, though changes in amplitude and duration can be seen as well. The spectrogram can be used to follow both types of spectral modulation (SM) that occur in speech: changes in the pattern of harmonics based on changes in fundamental frequency, and modulations of the spectral envelope as a result of vocal-tract filtering. Thus changes in the pitch of the voice can be seen (and measured to some extent) on narrow-band spectrograms, while broad- band spectrograms show well the configuration of formants and formant transitions, mirroring vocal-tract shapes and changes in shape over time.

Changes in speech-segment bandwidth are clearly suggested on the spectrogram. The range of segmental bandwidth varies from the exclusively low-frequency periodic energy of nasalization (low-pass at 300 Hz), to the broad-band noise of /sh/ to the many-frequency click of a /t/ or a /k/. The overall bandwidth of speech sounds, characterizing the spectral window in which a listener must be ready to listen for the next speech segment, is at least 6000 Hz, ranging from the lowest voice-pitch fundamental (about 100 Hz) to the high frequencies included in speech noises of clicks and frication.

The nature of a sound's spectrum is a principal indicator of its apparent quality or timbre. The timbre of a sustained sound (as opposed to a transient one) is a function of both the source characteristic—whether complex tone or noise— and the spectrum envelope, that is, the distribution of energy across frequency. While the fundamental frequency of the voice yields a perception of voice pitch, which changes over time, and

changes in overall amplitude yield structure of successive events in a sequence, it is the timbre that is most associated with differences among the successive elements.

The spectrogram reveals that connecting all segments in speech, whether they involve smaller or greater changes in bandwidth, are gradual spectral changes (the formant transitions). These "transitional segments" (as Fant [21] would call them), sometimes connect two equal-amplitude, same-source segments (as in a sequence of two vowels, /a-i/), sometimes connect segments that are similar in source but different in amplitude (e.g., from vowel to "semivowel" consonant such as /m, l, w/), and can also appear as spectral bridges between segments that differ in both source and amplitude, such as the periodic transitions between /s/ and /a/, or the noisy transitions between /k/ and /a/.

*Duration and speech sequences*

What is the time base according to which all these changes in amplitude and spectrum occur? In order to consider all aspects of narrative speech (not just the signal of isolated monosyllable or one-word utterances), it is necessary to define two time bases. There is one set of durations involved in sequences comprising phonemes. These durations range from a few milliseconds (e.g., 5 msec for a /p/ burst) to more than 100 ms (e.g., the sequence of segments making up an unvoiced affricate may take 150 msec to complete).

On the other hand, changes in the signal that pertain to prosody such as intonational contours in fundamental frequency, fluctuations of higher- and lower-amplitude vowel peaks as stress cues, and sequences of longer and shorter syllables, take place over many milliseconds to seconds. A short sequence of only two vowel peaks can establish a unit of rhythm (weak-strong = iambic, strong-weak = trochaic, etc.), but the rhythms of narrative speech are established over many seconds, from phrase to phrase and sentence to sentence. Students of sequence perception per se, such as Jones [29, 30] and Martin [40], have suggested rules for the perception of rhythms that include element durations as short as a few milliseconds, and patterns that emerge over long periods of time.

Perhaps the most interesting durational aspect in the configuring of the speech signal is that these two time bases can be shown to interact. For example, as noted by Lehiste [37], in speech that is more rapid than normal, prosodic changes come about more quickly (rhythmic units occur in shorter times), and phoneme-segment durations are shortened. Our own measurements show that this effect of speeding up speech can be seen in all types of speech segments: consonant closures and noises become shorter, semi-vowels are shorter, vowels remain at their high-amplitude peaks for less time, and formant transitions bridging jumps in bandwidth or amplitude or spectrum occur more quickly, with the rate of change increased. Such changes are significant, for the formant transitions of /w/ recorded at one (computer-generated) speaking rate can be heard as /b/ when inserted into a sentence produced at a slower rate (described in Hillenbrand et al. [24]). These authors suggest that listeners can " 'take into account' overall speaking rate in making ... decisions on phonemic contrasts" (p. 163).

Slowing speech from a normal rate causes lengthening primarily in the high-amplitude portions of vowels. However, consonants can be lengthened as well if the need arises. For example, our measurements show that if a talker is trying to achieve extremely clear articulation, as for a hard-of-hearing listener or a child, he/she may lengthen consonant noises, extend the low-amplitude portions of semivowels, and exaggerate the closure interval in stops and affricates.

**Psychoacoustics of nonspeech temporal patterns**

*Design of nonspeech auditory sequences*

If speech comprises sequences of different acoustic segments, students of speech perception should be interested in the perception of sequences of all types of auditory events. Most of the psychoacoustic literature deals primarily with isolated sounds that do not change over time (tones, noise bursts, and combinations of the two). However, there are many experiments on sequential arrays whose results are relevant to the speech signal.

By definition, this "psychoacoustics of sequences" includes any study of how listeners perceive sounds that change over time. Keeping our list of speech-signal properties in mind, we can classify the time-varying sounds that have been studied in nonspeech psychoacoustics in terms of the dimensions used to articulate auditory changes over time (cf. Hirsh [26]; In press). Such a classification can then serve as a vantage point from which to compare the auditory capabilities that have been demonstrated for nonspeech time-varying sounds, and the abilities that seem to be required by speech sequences.

Any sequence of individual sound events may be characterized as a succession of elements that differ from each other in amplitude only, spectrum only, duration only, or combinations of such attributes. In psychoacoustic studies, measurements of duration must distinguish between duration of individual elements and the timing of element onsets. Note that element duration cannot be studied in the absence of changes in other dimensions—experiments in "duration discrimination" necessarily make use of sequences involving either amplitude or spectral changes to define the elements whose durations are compared. For example, two or more contiguous sounds must be marked by differences in amplitude or spectrum before durations can be judged. On the other hand, sequences made of noncontiguous sounds, including those used in judgements of timing between element onsets, involve points of 100% amplitude modulation, from elements to intervening silences.

*Experiments on auditory sequence perception*

*Amplitude-modulated (AM) sequences*

Relevant psychoacoustical experiments here include those on detection and discrimination of amplitude modulation of tones or noises, auditory fusion of spectrally identical elements, gap detection (i.e., detection of 100% modulation), temporal masking where the signal and masker are identical in spectrum, duration discrimination and temporal pattern recognition where sequential elements are spectrally the same (as in Morse code sequences), and intensity streaming. (For a

survey of experiments on these and other psychoacoustical questions, cf. Moore [43].)

Results of such experiments are relevant to both segmental and suprasegmental aspects of speech. A number of segmental cues involve amplitude fluctuations: from vowel "ceiling" to consonants (different modulation depths may act as cues to different consonant classes), from consonant to consonant within clusters (e.g., from /f/ to /s/), from point to point within a consonant (e.g., from burst of /k/ to aspiration noise). Nonspeech experiments may tell us whether such changes can be perceived (given their magnitude and time course) and how neighboring segments may influence each other, particularly in terms of masking.

As Cole and Scott [11] have noted, temporal changes in amplitude are associated with suprasegmental aspects of speech, as well: shifts from sound to pause, amplitude fluctuations signalling stress contours, the changes in waveform envelope associated with speech rhythms. Nonspeech experiments may help us to make predictions as to the efficacy of such changes as perceivable cues to speech prosody.

For example, to study *amplitude modulation*, Patterson et al. [51] used AM noise and found that listeners' ability to detect differences in the depth of modulation varied with modulation rate and noise bandwidth. The faster the rate, the greater the depth needed to hear the changes; the wider the bandwidth of the noise, the smaller the depth needed. Thus it is possible in speech perception that listeners who are using a 6000-kHz window, to include the noises of speech, may be able to hear even the smallest changes in amplitude envelope that are correlated with changes in speech sounds. But speech is not all noise; what of AM tones? Riesz [53] found that listeners were most sensitive to AM tones at a rate of 2–4 Hz, or one change every 250 msec. Is the 2–4 Hz modulation rate important for the perception of syllable succession (perception of speech rhythms), and a faster modulation rate more related to differentiating phonemes? Of course these experiments can only be suggestive with relation to rules for speech perception. There are many steps between the sounds tested here and speech sounds.

Experiments in *auditory fusion* (cf. Hirsh [25] for an early review) have found that two same-spectrum elements can be heard out at onset asymmetries as small as 2 msec. This is certainly well below the temporal threshold required by the structure of speech sounds. Bilabial bursts, the shortest segments in English, seldom last less than 2 msec. Whether such short, low-amplitude events are masked by successive formant transitions remains to be determined, but certainly a listener should be able to discriminate the difference between a sound preceded by the burst and one without it. Instances of "overlapping" sounds in speech, such as onset of nasal formants during a vowel or consonant preceding a nasal conson-ant, may take use of such abilities, though they do not stretch the demonstrated limits. Such over-lapping characteristically leads (and follows) the actual nasal consonant waveform portion by some tens of msec.

In experiments in *gap detection*, one can test whether interpolated gaps can be heard (a kind of fusion question), or measure directly a lis-tener's ability to discriminate changes in gap dura-tions. A number of experiments (e.g., Abel [1, 2]; Divenyi and Danner [16]; Chistovich [8]) have found that listeners can discriminate 10% changes in either filled or unfilled temporal inter-vals. Such abilities may be important for hearing out differences in VOT ("filled intervals") of stops and affricates. Surveys of VOT steps in natural speech (cf. Zue [63]) indicate that within-talker differences (e.g., from /b/ to /d/ to /g/) are usually on the order of 25–50%, well above the limits demonstrated in nonspeech experiments. Gap-detection abilities are not taxed in percep-tion of any stop/affricate closure, which even in rapid speech are several tens of msec. long.

*Temporal masking* with same-spectrum sounds (tone-on-tone, noise-on-noise) has been studied recently in measurements of psychophysical tun-ing curves (e.g., Moore et al. [44]). These results may be relevant to the interactions between speech segments that are similar in spectrum, such as the burst and following higher-amplitude frication of affricates. Observations of the relative levels of these two speech-noise segments, com-pared with the psychoacoustic noise-on-noise re-sults, suggests that listeners can in fact hear the

preceding burst—though again we must re-member that the speech segments differ in both amplitude and spectrum. In order to predict more accurately rules for perception of the speech se-quences, we must consider masking experiments where masker and target are distinct in both amplitude and spectrum (see below).

## Spectrally-modulated (SM) sequences

Experiments using sounds of this type include studies of frequency modulation of continuous pure tones, pure-tone glide perception where glides are presented in sequence with other ele-ments (such as CF tones), complex-tone modula-tion (nonspeech formant transitions), timbral changes (e.g., discrimination of noises with re-gard to central frequency and/or bandwidth, dis-crimination of complex tones differing in spec-trum), temporal masking where signal and masker differ in spectrum, and temporal-order studies of two or more elements differing in spec-trum (tones of two frequencies, click-tone, etc.). Note that we will not consider under this category any experiments where sounds are not contigu-ous. As explained above, sequences of elements separated by silences properly should be thought of as sequences that include points of 100% amplitude modulation: this is potentially an im-portant distinction for interpreting results with re-ference to speech perception.

Experiments using spectrally-modulated sounds are relevant to segmental aspects of speech with regard to cues based on changes in formant frequencies and timbre, and questions of masking by different-spectrum contiguous seg-ments (e.g., voiced formant transitions masking preceding /b/ burst)—though again, the speech segments obviously differ along a number of di-mensions, not just spectrum. One suprasegmental parameter involving spectral changes is the tem-poral pattern of fundamental frequency, which may be relevant to speaker identification and is certainly important for the perception of stress and rhythm.

We might note here that although we have in-cluded both pitch differences and timbre differ-ences under "spectral modulation", it may turn out that perceptually these two aspects of spec-trum are distinct. We are not invoking the

dichotomy between "spectral pitch" and "virtual pitch" (or their terminological variants), but defining two larger categories of sound properties. For example, the "pitch" of a given speech segment and its identity as either periodic or aperiodic may be equally important for speech perception.

In fact, as we have suggested elsewhere (Lauter [31]), this distinction between pitch and timbre may provide a basic distinction between speech and music. It is quite usual in speech that successive elements are very different in timbre, e.g., a complex-tone vowel is succeeded by a noisy unvoiced fricative. Very rapid shifts in timbre can occur due to changes in sound source, of spectrum envelope, or of both. In contrast, except for experiments by a few composers, e.g., Webern, Schoenberg, Ligeti, in "klangfarbenmelodie", and some modern compositions using synthesizers, musical melodies typically comprise sequences of elements that may change in pitch, amplitude, and duration, but do not vary in timbre. I.e., one does not usually listen to melodies where one note is played on a flute, the second on a drum, the third on a gong, the fourth on a violin, etc. Such constructions may be used for musical humor (cf. some of Spike Jones' routines), but are not the rule in classical composition. There is evidence (see below) suggesting that constant-timbre and timbre-modulated sequences are perceived in basically different ways, which may point to a psychoacoustical distinction between sounds that are more like speech and sounds more like music.

A number of experiments have examined listeners' abilities to detect *frequency modulation* of pure tones (e.g., Shower and Biddulph [55]), and how performance changes as a function of rate and depth of modulation. Similar results have not been gathered for shaped complex tones (with formant structure), either for pitch change or timbre change, although this is clearly relevant to both speech and music perception.

*Discrimination or identification* tasks using sequences of spectrally-different elements represent tests of sequence perception. Pure tones have been studied in this way, to determine difference limens for frequency. Noises also have been tested, to determine difference limens for center frequency and bandwidth; and complex tones have been used to observe how spectral pitch and virtual pitch are compared by listeners. Distinctions that listeners can make between the elements of such sequences suggest how speech segments may fare when neighboring segments are different in spectrum (whether it is pitch or timbre that is the focus of the difference).

*Temporal masking* using sounds that differ in spectrum is one of the most common types of psychoacoustical experiments. However, most masking experiments also contrast the amplitude of the signal and masker, and most temporal masking experiments consider a range of separations between the two elements. Thus these experiments are properly AM + SM, and will be considered in the next section. As far as we know, there are no data for combinations of different-spectrum, equal-amplitude target and masker, with no gap between. This is perhaps not very serious for speech perception interests, since most contiguous speech segments that differ radically in spectral envelope (i.e., have different sources) are different in amplitude as well.

Finally, listeners have been asked to discriminate or identify sequences composed of *temporally overlapping* tones of two frequencies, or overlapping noise and tone. Listeners' abillities to discriminate such two-element combination sounds were studied by Efron [18] and Patterson and Green [49]; both experiments found that with only 1 to 2 msec onset asymmetry, listeners could perceive quality differences (cf. Green [22]) that helped them tell the difference between the two orders. However, to identify the order as AB versus BA, Hirsh [25] found that listeners needed a minimum of 17 msec onset asymmetry; this rule obtained whether the two elements were a high tone and a low tone, or a noise and a tone. These results may have some relevance for the identification of voicing in initial plosives (e.g., Pastore [48] or of voiced fricatives. Though listeners may simply learn the characteristic "combined" (laryngeal plus noisy) spectrum of voiced fricatives, one might speculate that sequential perception is important as well. The waveform of voiced fricatives shows that for some talkers, in the early segments (e.g., the first 10–30 msec) of these consonants, voicing predominates, with the noise

gaining prominence only toward the second half of the sound.

In a series of experiments directed to determining how listeners perceived *melodies* of three contiguous tones, Divenyi and Hirsh [13, 14, 15] manipulated such variables as frequency range of the pattern, frequency ratios between tones, duration of tones, and effect of a fourth tone. Among their results was the observation that frequency range interacted with tone duration to affect identification performance: melodies spanning a frequency range of about one octave could be identified with tones of only 2-7 msec durations; however, restricting the range within one octave, to less than one-third octave, worsened identification performance at short durations. Thus it appears that the spectrum of very short sequential events can be identified with a high degree of accuracy; this may be of relevance to the perception of stop bursts, which may be as short as 5 msec. However, a counteracting influence in the speech case (the spectral proximity of neighboring events) is also suggested by these experiments: is the perception of a stop burst actually made more difficult because the succeeding sound (whether voiced formant transitions as in voiced stops, transitions in aspiration noise as in voiceless stops, or higher-amplitude frication in affricates) has a spectrum initially "continuous" with that of the burst?

Divenyi and Hirsh also found that identification was better for "unidirectional" frequency patterns (low-middle-high, high-middle-low) than for sequences that changed direction in mid-pattern. This last result may augur well for the usefulness of formant transitions in speech, given the typical durations of unidirectional transitions. Although formant transitions do change direction, they typically move from higher to lower (or v.v.) for several tens of msec before changing; such a change in direction is heard as a change in place of articulation, whether the change is from vowel to vowel, consonant to consonant, or between a vowel and a consonant.

Of course formant transitions acoustically are more like *frequency glides* than series of frequency steps. The perception of gliding pure tones in sequence with steady-state frequencies has been studied in a number of experiments. For

example, a series of papers by Nabelek and Hirsh [45], and Nabelek, Nabelek and Hirsh [46, 47] reported listeners' abilities to discriminate. the rate of a glide as a function of frequency excursion of the glide, glide duration, and duration of steady-state framing frequencies. One result was that "optimum glide rates" (i.e., change in frequency per unit time) involved very similar "optimum glide durations"—20 to 28 msec—over a range of frequencies from 250 to 4 kHz. The authors noted the relevance of these findings to speech: "It is interesting to see that these optimal transition durations do not depend on frequency region and are close to the durations of transitions found ... to be important for the discrimination of [synthetic] speech sounds." They concluded that perception of this characteristic in speech sounds may represent only one instance of a general psychoacoustical capability: "these results indicate that the best discriminability of large delta-f for transition durations around 30 msec is a general property of hearing and that it does not appear only in connection with speech sounds" (p. 1518).

Other experiments have examined *interactions* of spectrum and duration. For example, subjects have been asked to match the pitch of a fixed tone to a variable one, with manipulations by the experimenter in the duration of the fixed tone. Liang and Chistovich [38] and Henning [23] found an interaction between tone duration and frequency difference limen: difference limens increase as the tones become shorter than 100 msec. In a study focused on the combined contribution of spectral and durational cues to the perception of two-tone sequences, Espinoza-Varas [20] presented listeners with two tone bursts separated by a 60-msec silence. He varied the duration of the two tones and their frequency, asking listeners simply to judge whether the two were (in any way) "same or different". Results indicated that even when the durational difference between the two was (somewhat) smaller than the duration DL and the frequency difference was (somewhat) smaller than the frequency DL, listeners could accurately discriminate between the two. Espinoza-Varas interpreted this demonstration of the perceptual "additive" quality of physical parameters of sounds in sequence as possibly re-

levant to "integration of multidimensional cues" in speech. In his discussion, the author cited an observation by Best et al. [4], that in the contrast between "stay" and "say", the discriminability of a difference in the duration of a silence interpolated between /s/ and the vowel can be reduced or eliminated by the introduction of additional differences in the formant transitions into the vowel. Espinoza-Varas concluded that "integration [of cues may be] a general capability that operates with both speech and nonspeech sounds" (p. 1693).

*Sequences with both AM and SM*

A variety of experiments can be characterized as using sounds that change both in spectral characteristics and amplitude over time. The most common are temporal masking paradigms. Experiments on discrimination and/or identification of sequences of noncontiguous sounds, such as melodic patterns or Morse-code-type rhythms, can also be classified in this way, as can a number of nonspeech studies where combinations of AM/ SM sounds have been used to study other phenomena, such as order identification of contiguous sounds, streaming, categorical perception of nonspeech sequential sounds, and discrimination of properties of sounds in context.

Relevance of such experiments to problems of speech perception include segmental distinctions involving concomitant changes in both amplitude and spectrum (e.g., the burst and succeeding frication of /ch/ differ in both amplitude and spectrum). This is of course the most common case in speech-sound construction, and thus nonspeech sequences using this design bear a closer resemblance to speech sounds than do sequences where amplitude or spectrum change exclusively. Among the related suprasegmental speech parameters are cues to stress and rhythmic patterns based on co-occurring changes in amplitude and fundamental frequency.

The most usual type of *temporal masking* experiment presents a pure tone as target preceded (forward masking sequence) or followed (backward masking sequence) by a noise burst. There is usually more backward than forward masking, and the "masked threshold" for the tone depends on relative levels of target and masker, durations,

and spectral resemblance (for a review, cf. Patterson and Green [50]. Thus in speech one might expect, for example, that because of formant transitions' "spectral bridges", succeeding speech segments have more chance to mask each other than if there were sudden spectral shifts from segment to segment. It is possible that other characteristics of formant transitions, such as gradual instead of abrupt changes in amplitude, may counteract the potential masking effect of spectral similarity between neighboring sounds.

Pollack [52] studied the combined effects of a preceding and following noise on a tonal target, with a sequence of three 50-msec sounds, separated by variable gaps. By manipulating the two gaps and testing detection of the tone, he found that in this combined masking situation there was an asymmetry of masker configuration. Specifically, in the presence of two maskers, varying the backward-masking interval had more effect on the masking shown at long forward-masking intervals than on the shorter intervals. Details of such interaction of sounds preceding and following a target may be important for the perception of stops and affricates, where the closure interval (analogous to Pollack's pre-target gap) in any given instance is longer than the associated VOT (analogous to Pollack's post-target interval preceding the higher-amplitude [vowel] "backward masker").

More elaborate sequences, constructed to resemble musical patterns, have been studied in a series of experiments on *"streaming"* (e.g., Miller and Heise [41], Bregman and Campbell [6], Bregman and Dannenbring [7], Bregman [5]). These experiments illustrate interactions among sequence variables such as the rate of modulation, the amount of spectral change, and the amount of amplitude change: according to the rate of change, sounds can be shown to "stream" based on either frequency or amplitude. Other experiments suggest that more gross spectral differences, such as differences in source (e.g., McAdams and Bregman [39], Wessel [62]) or in ear-of-presentation (e.g., Efron and Yund [19], Deutsch [12]) can provide the basis for streaming as well.

Consideration of the results of the monaural streaming experiments, particularly with regard

to the importance of rate of SM/AM for "stream segregation," lead one to wonder why the rapidly changing segments of speech do not undergo streaming. Cole and Scott [11] suggested a demonstration showing that in fact streaming principles do hold for speech sounds—when they are cycled (played repeatedly): for example, after five to six repetitions, /ta/ streams into a short "s" followed by "da".

These authors suggested that the rapidly modulated sequence of speech segments resists stream segregation because of the presence of formant transitions, which provide SM/AM bridges between the extremes of spectral and amplitude values. Bregman and Dannenbring [7] tested a similar hypothesis, by examining streaming in tone sequences with and without interpolated pure-tone glides. They found that in fact frequency separation and rate of change could be increased beyond the normal streaming threshold when glides connected successive tones.

In a study using sounds from *different sources*, Miller et al. [42] prepared a sequence of a bandpass noise and a square wave, with the noise leading or lagging tone onset over a range from −10 to +80 msec. In separate tests subjects were asked to discriminate or label the sequences as "noise" or "no noise". The authors summarized their results as showing that "discrimination was best across a noise-lead-time boundary of about 16 msec, where labeling also shifted abruptly," and pointed out that these results were "highly similar to those reported for the categorical perception of synthetic plosive consonants differing in voice onset time" (p. 410).

Warren and colleagues (for a review, see Warren [58]) also studied sounds involving changes in source over time. In one experiment they constructed sequences using a square wave, an octave-band noise, a pure tone, and an excerpt of a real /i/. This "unit sequence" was then presented in a continuously cycling mode, i.e., with no intervals between successive sets of four elements. Listeners were asked to report the temporal order of the four elements, as element duration was changed. Results showed that listeners could not identify the temporal order of these four cycling sounds until each element was at least 650 msec long. The authors pointed out the obvious similar-

ity between the acoustical profile of their sequences and speech sequences, and the obvious disparity between the element duration required in their test and that existing in speech. As they and others have suggested, it is possible the disparity between their results and those seen in speech arose from two basic dissimilarities between the test sequences and speech sequences: 1) the Warren et al. patterns were continually recycled; sequence identification was not tested with the sequences in isolation (Hirsh [27]); and 2) the elements changed abruptly from one to the next; they lacked the gradual SM/AM changes that are present in speech.

Follow-up experiments (cf. Thomas et al. [57]; Cole and Scott [11]; and Dorman et al. [17]) have demonstrated that the ordering task can be done with shorter element duration if all elements are vowels (i.e., no source changes over time), with even shorter durations if there are silences between the vowels, and at speech rates if transitions connect the vowels. Experiments in our laboratory (cf. Lauter [31]) indicate that subjects can learn to identify temporal order characteristics of multi-source sequences, with element durations in the speech range (e.g., 25–100 msec), even without transitions, as long as the sequences are presented in isolation, i.e., not cycled. Thus as Cole and Scott [11] and others have suggested, the original Warren et al. [59] results are probably not surprising in light of the basic differences between the nonspeech perceptual task and speech, and the original results should not be taken as contradictory to our suppositions as to the rules that pertain in speech perception. In fact, differences in performance correlated with differences in task may help us to understand more about some of the "essential" characteristics of speech sequences.

Another series of studies on *ten-tone sequences* has been reported by Watson and colleagues [56, 60, 61]. Although these experiments are not directed to sequence perception per se, they provide information as to how well listeners can resolve individual elements within a sequence. Watson et al. used sequences of 10–40 msec contiguous pure tones, varied a number of tone parameters such as frequency and intensity, and studied the effect on sequence discrimination of changes

in these parameters as a function of tone position in the sequence and stimulus uncertainty. Under conditions of low stimulus uncertainty (i.e., when listeners knew which tone was the target for a possible change), frequency and intensity discrimination were as good as for isolated (non-sequenced) tones. These experiments are certainly related to speech perception, even though speech sequences rarely consist of trains of same-timbre elements. The relevant connections involve not only the values of temporal resolution tested, but also questions of selective attention to individual segments within sequences. The 40 msec minimum duration of Watson's tones is an average value for speech segments (the reason for choosing it). As an average, it perhaps underestimates the limits of resolving power for any one element, but over a 400 msec sequence, may provide a fairer reflection of how closely one can listen to speech segments, than auditory flutter fusion experiments.

Also, in terms of listening within sequences for acoustic changes, these findings may have relevance to the ways in which the speech signal is "tailored" for accurate listening. Though it may seem that listeners faced with novel narrative speech cannot know when to attend for important distinctions in speech sounds, Jones [29, 30] and others have suggested that in fact speakers use prosodic patterns to help the listener predict when to listen for important acoustic distinctions, by placing acoustical high-information points at peaks of rhythmic emphasis. In a recent paper, Leek and Watson [36] addressed the question of how listeners learn to listen within such sequences for potential changes. They measured how long each of a group of listeners took to learn "when to listen" (i.e., to which element in the sequence of ten) in a ten-tone sequence for potential level changes. Then the same listeners were presented with a new set of sequences; the experimenters found the subjects learned "when to listen" much more quickly on the new, but related, task. The authors guessed that their observations of the listeners' capabilities may be related to speech perception, where listeners must acquire the knack of listening within sequences for small changes that could cue, e.g., the difference between /t/ and /s/. One could also combine Jones' sugges-

tions about rhythmic pointers in speech, and the ten-tone configuration, to study whether listeners can learn to use "prosodic" aspects of sequences to find out "when to listen" for important events. For example, one could present listeners with four sequences: the first two with amplitude highlighting a target element, and the second two a discrimination pair, with a small change in frequency in the element highlighted in the first two sequences. Learning curves could then be compared to see whether listeners learned the discrimination task more quickly with the rhythmic cueing than without it.

## Speech-perception psychoacoustics

Most studies of sound parameters important for speech perception have been examined in sounds that more or less resemble speech sounds, but that can be identified as speech sounds. The majority of these experiments have used different sorts of "synthetic speech", presented almost always as isolated elements (syllables to words). Fewer experiments have made use of real speech, edited or transformed in a variety of ways (cf. Lauter [35]), but still identifiable in terms of actual speech sounds.

### Psychoacoustics of speech-like sequences

As the foregoing review implies, there is possibly much to be found out about human auditory perception that we cannot discover if subjects are forced to label everything they hear as one of forty phonemes, or combinations thereof. Sound sequences designed according to the principles of speech-segment construction may hold the potential of telling us both about basic and more sophisticated capabilities of the auditory system that are evoked in the perceptual tasks related to speech and music.

For example, sequences may contain from two to many events, with parameters of amplitude, spectrum (pitch and timbre may be manipulated separately), and duration held constant or changed in different ways over time. Measures taken from close analysis of speech segments can serve as guidelines for the design of such

nonspeech sequences, so that perception of sequences more unlike speech can be compared with performance on sequences that are acoustically very like speech (but would not be identifiable as speech). One *example of a stimulus paradigm* sampling such extremes would be sets of three-element patterns, one set (most unlike speech) made with contiguous pure tones, with tones differing only in frequency, and tone duration longer than 200 msec. A more speech-like version of such sequences would shorten the tones to phoneme-like durations (e.g., Watson's 40 msec or shorter), or use complex tones instead of pure tones. A configuration most like speech would use three elements each of which differed in amplitude, pitch, timbre, and duration, with the changes from element to element along each of these dimensions similar to those found from segment to segment in speech. Thus one could create a nonspeech sequence acoustically resembling, e.g., /t/ that no listener would hear as /t/ (cf. Lauter [31]).

The variety and range of psychoacoustical tests that could be pursued with such stimuli are obvious, and could include study not only of perception of single cues in single elements (such as in Watson's work), but also of the interaction of cues both *within the same element* and *across the sequence* (such as in the study reported by Espinoza-Varas). Elaborations of such sequences would include not only series where changes in all parameters occur at the same times, but also (more like some aspects of speech) sequences where one parameter changes at one time, and another changes at another time (as in nasalization re formant transitions, or *F*0 fluctuations re phonemic segmentation).

*"Comparative psychoacoustics": speech v::.
nonspeech*

The eventual ideal of such testing would be to find some way to compare performance in the same listeners on a series of nonspeech sounds and related speech sounds. Past instances of this approach have involved starting with phenomena first observed in speech perception, such as categorical perception or ear advantages, and looking to see if the same phenomena can be demonstrated with nonspeech sounds. For example, the experiment using sequences of a noise and a complex tone reported by Miller et al. [42], was designed to see whether nonspeech sequences could evoke categorical perception similar to that seen for stop consonants. Although performance in the same listeners on speech and nonspeech sounds was not tested, the authors compared the nonspeech discrimination and identification performance with results on similar tasks using speech sounds, as reported in the literature. Based on such comparisons, the authors were able to conclude that "categorical perception of sounds is not unique to speech ... and may be a general property of sensory behavior."

Another type of "comparative psychoacoustics" has made use of ear advantages as a dependent variable common to a variety of test sounds (cf. Berlin and McNeill [3]). For example. Lauter [31, 32] reported results of presenting subjects with sets of dichotic sounds, and asking them to label the sounds in only one ear from trial to trial. The dependent variable for all sounds was ear advantage (EA), computed by acquiring an overall score for performance in the left ear, a score for the right ear, and subtracting them. Several sets of sounds were tested, ranging from pure-tone melodies involving changes only in frequency from tone to tone, and 80 msec between tone onsets. to faster melodies. to melodies made with noise bands, to sequences involving multidimensional changes in elements. to real speech vowels and stop CVs to synthetic stop CVs. Each listener was tested on all sounds. Results indicated that sequences that acoustically were least like speech evoked ear advantages that were very different from those seen for the speech sounds, while sequences that were designed to be more like speech (e.g., with multidimensional changes over time) evoked ear advantages similar to speech EAs. In a subsequent review (Lauter [33]), results from a variety of experiments in dichotic listening were analyzed to show that a few sequential properties seem to be important for evoking EAs more or less like those seen for speech sounds. Among these were: rate of events (or duration of contiguous elements), element bandwidth, and the number of dimensions changing over time. Future dichotic work with se-

quences that are designed to mimic different types of speech sounds (e.g., vowels, plosives, prosodic changes) may lead to insights regarding the auditory abilities used to perceive these speech distinctions.

Sounds that are not heard as speech (or music) used to study basic auditory perceptual abilities provide two advantages for the student of auditory perception. First, they avoid problems of familiarity and overlearning associated with the everyday sound systems of speech and melody. Second, they provide all the power of modern psychophysics, with complementary aspects of control over stimulus and response. As the previous review implied, the use of such sounds makes a vast array of earlier psychoacoustic observations available as potential guidelines for designing sequences, and for interpreting experimental results.

It is probable that if we go on requiring listeners to label test sounds as speech sounds, we will be limited in how much we can find out about how the auditory system works to achieve speech perception. Certainly if we hope to discover anything about the nervous-system mechanisms underlying behavioral auditory performance, we need to systematically study a range of stimuli and tasks. As we have suggested elsewhere (Lauter [34]), it is possible that nervous-system complexity (from periphery, into successive levels of the CNS) is in some way correlated with the complexity of the sensory-perceptual situation. In order to be able to study these hierarchies of test and perceptual system, we need to approach the design of psychoacoustical test sounds in a systematic, analytical way. This involves first defining the extremes of complexity, and then a selection of the dimensions that will allow us to measure performance along a continuum ranging from simple to complex.

## References

[1] S.M. Abel, "Duration discrimination of noise and tone bursts," *J. Acoust. Soc. Amer.* 51, 1219–1223, (1972).
[2] S.M. Abel, "Discrimination of temporal gaps." *J. Acoust. Soc. Amer.* 52, 519–524, (1972).
[3] C.I. Berlin and M.R. McNeil, "Dichotic listening," In:

N.J. Lass (Ed), *Contemporary Issues in Experimental Phonetics*, NY: Academic, pp. 327–388. (1976).
[4] C.T. Best, B. Morrongiello, R. Robson, "Perceptual equivalence of acoustic cues in speech and nonspeech perception," *Percept Psychophys.* 29, 191–211, (1981).
[5] A.S. Bregman, "The formation of auditory streams," In: J. Requin (Ed), *Attention and Performance*, Vol. 7. NJ: Lawrence Erlbaum Assoc. (1978).
[6] A.S. Bregman and J. Campbell, "Primary auditory stream segregation and perception of order in rapid sequences." *J. Exp. Psych.* 89. 244–249, (1971).
[7] A.S. Bregman and G.L. Dannenbring, "The effect of continuity on auditory stream segregation," *Percep. Psychophys.* 13, 300–312, (1973).
[8] L.A. Chistovich, "Discrimination of the time interval between two short acoustic pulses." *Soviet Physics: Acoustics* 5, 493–496, (1959).
[9] L.A. Chistovich, N.A. Fyodorova, D.M. Lissenko, M.G. Zhukova, "Auditory segmentation of acoustic flow and its possible role in speech processing." In: G. Fant and M.A.A. Tatham (Eds), *Auditory Analysis and Perception of Speech*, London: Academic. pp. 221–232. (1975).
[10] R.A. Cole and B. Scott, "Perception of temporal order in speech: The role of vowel transitions," *Can. J. Psych.* 27, 441–449, (1973).
[11] R.A. Cole and B. Scott, "Toward a theory of speech perception," *Psych. Rev.* 81, 348–374, (1979).
[12] D. Deutsch, "Two-channel listening to musical scales," *J. Acoust. Soc. Amer.* 57, 1156–68, (1975).
[13] P.L. Divenyi and I.J. Hirsh, "Identification of temporal order in three-tone sequences," *J. Acoust. Soc. Amer.* 56, 144–151, (1974).
[14] P.L. Divenyi and I.J. Hirsh, "The effect of blanking on the identification of temporal order in three-tone sequences," *Percep. Psychophys.* 17, 246–252, (1975).
[15] P.L. Divenyi and I.J. Hirsh, "Some figural properties of auditory patterns." *J. Acoust. Soc. Amer.* 64, 1369–1385. (1978).
[16] P.L. Divenyi and W.F. Danner, "Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra," *Percep. Psychophys.* 21, 125–42. (1977).
[17] M.F. Dorman, J.E. Cutting, I.J. Raphael, "Perception of temporal order in vowel sequences with and without formant transitions," *J. Exp. Psych.: Hum. Percep. Perf.* 184, 121–129, (1975).
[18] R. Efron, "Conservation of temporal information by perceptual systems," *Percep. Psychophys.* 14, 518–530, (1973).
[19] R. Efron and E.W. Yund, "Ear dominance and intensity independence in the perception of dichotic chords," *J. Acoust. Soc. Amer.* 59, 889–98, (1974).
[20] B. Espinoza-Varas, "Integration of spectral and temporal cues in discrimination of nonspeech sounds: A psychoacoustic analysis," *J. Acoust. Soc. Amer.* 74, 1687–1694, (1983).
[21] G. Fant, "Descriptive analysis of the acoustic aspects of

speech," *Logos 5: 3–17. Reprinted in: G. Fant (1973), Speech Sounds and Features.* Cambridge: MIT Press, pp. 17–31, (1962).

[22] D.M. Green, "Temporal acuity as a function of frequency," *J. Acoust. Soc. Amer.* 54, 373–379, (1973).

[23] G.B. Henning, "A comparison of the effects of signal duration on frequency and amplitude discrimination," In: R. Plomp and G. Smoorenburg (Eds), *Frequency Analysis and Periodicity Detection in Hearing,* Leiden, Netherlands: Sijthoff, pp. 350–359, (1971).

[24] J. Hillenbrand, F.D. Minifie, T.J. Edwards, "Tempo of spectrum change as a cue in speech-sound discrimination by infants," *J. Speech Hearing Res.* 22, 147–165, (1979).

[25] I.J. Hirsh, "Auditory perception of temporal order," *J. Acoust. Soc. Amer.* 31, 759–767, (1959).

[26] I.J. Hirsh, "Temporal aspects of hearing," In: D.B. Tower (Ed), *The Nervous System,* Vol. 3, *Human communication and its disorders,* NY: Raven Press, pp. 157–162, (1975).

[27] I.J. Hirsh, "Order of events in three sensory modalities," In: S.K. Hirsh, D.H. Eldredge, I.J. Hirsh, S.R. Silverman (Eds), *Hearing and Davis: Essays honoring Hallowell Davis,* St. Louis: Washington University Press, pp. 159–173, (1976).

[28] I.J. Hirsh, "Auditory perception and speech," In: R.C. Atkinson, R. Hermstein, G. Lindzey and D.R. Luce (Eds), *Stevens' Handbook of Experimental Psychology,* NY: Wiley, (in press).

[29] M.R. Jones, "Time, our lost dimension: toward a new theory of perception, attention, and memory," *Psych. Rev.* 83, 323–355, (1976).

[30] M.R. Jones, "Auditory patterns: Studies in the perception of structure," In: E.C. Carterette and M.P. Friedman (Eds), *Handbook of Perception,* Vol. 8, *Perceptual Coding.* NY: Academic, pp. 255–288, (1978).

[31] J.L. Lauter, "Dichotic identification of complex sounds," Unpublished PhD dissertation, Washington University, St. Louis, (1979).

[32] J.L. Lauter, "Dichotic identification of complex sounds: absolute and relative ear advantages," *J. Acoust. Soc. Amer.* 71, 701–707, (1982).

[33] J.L. Lauter, "Stimulus characteristics and relative ear advantages: A new look at old data," *J. Acoust. Soc. Amer.* 74, 1–17, (1983).

[34] J.L. Lauter, "Auditory System," In: A.L. Pearlman and R.C. Collins (Eds), *Neurological Pathophysiology,* 3d. ed. NY: Oxford Univ. Press, pp. 86–109, (1984).

[35] J.L. Lauter, "The acoustic phonetics of real speech," (In preparation).

[36] M.R. Leek and C.S. Watson, "Learning to detect auditory pattern components," *J. Acoust. Soc. Amer.* 76, 1037–1044, (1984).

[37] I. Lehiste, *Suprasegmentals,* Cambridge: MIT Press, (1970).

[38] C. Liang and L.A. Chistovich, "Frequency difference limens as a function of tonal duration," *Soviet Physics: Acoustics* 6, 75–88, (1961).

[39] S.E. McAdams and A.S. Bregman, "Hearing musical streams," *Computer Music J.* 3, 26–43, (1979).

[40] J.G. Martin, "Rhythmic (hierarchical) versus serial structure in speech and other behavior," *Psych. Rev.* 79, 487–509, (1972).

[41] G.A. Miller and G.A. Heise, "The trill threshold," *J. Acoust. Soc. Amer.* 22, 637–638, (1950).

[42] J.D. Miller, C.C. Wier, R.E. Pastore, W.J. Kelly, P.J. Dooling, "Discrimination and labelling of noise-buzz sequences with varying noise-lead times: An example of categorical perception," *J. Acoust. Soc. Amer.* 60, 410–417, (1976).

[43] B.C.J. Moore, An introduction to the psychology of hearing, NY: Academic, (1982).

[44] B.C.J. Moore, B.R. Glasberg, B. Roberts, "Refining the measurement of psychophysical tuning curves," *J. Acoust. Soc. Amer.* 76, 1057–1066, (1984).

[45] I. Nabelek and I.J. Hirsh, "On the discrimination of frequency transitions," *J. Acoust. Soc. Amer.* 45, 1510–19, (1969).

[46] I.V. Nabelek, A.K. Nabelek, and I.J. Hirsh, "Pitch of tone bursts of changing frequency," *J. Acoust. Soc. Amer.* 48, 536–553, (1970).

[47] I.V. Nabelek, A.K. Nabelek, and I.J. Hirsh, "Pitch of sound bursts with continuous or discontinuous change of frequency," *J. Acoust. Soc. Amer.* 53, 1305–1312, (1973).

[48] R.E. Pastore, "Categorical perception: a critical reevaluation," In: S.K. Hirsh, D.H. Eldredge, I.J. Hirsh, S.R. Silverman (Eds), *Hearing and Davis: Essays honoring Hallowell Davis,* St. Louis: Washington University Press, pp. 253–264, (1976).

[49] J.H. Patterson and D.M. Green, "Discrimination of transient signals having identical energy spectra," *J. Acout. Soc. Amer.* 48, 894–905, (1970).

[50] R.D. Patterson and D.M. Green, "Auditory masking," In: E.C. Carterette and M.P. Friedman (Eds), *Handbook of Perception,* Vol. IV, *Hearing,* NY: Academic, pp. 337–361, (1978).

[51] R.D. Patterson, D. Johnson-Davies, R. Milroy, "Amplitude-modulated noise: The detection of modulation versus the detection of modulation rate," *J. Acoust. Soc. Amer.* 63, 1904–1911, (1978).

[52] I. Pollack, "Interaction of forward and backward masking," *J. Aud. Res.* 4, 63–67, (1964).

[53] R.R. Riesz, "Differential intensity sensitivity of the ear for pure tones," *Phys. Rev.* 31 (series 2), 867–875, (1928).

[54] E.W. Scripture, *The Elements of Experimental Phonetics,* NY: Scribners, (1902).

[55] E.G. Shower and R. Biddulph, "Differential pitch sensitivity of the ear," *J. Acoust. Soc. Amer.* 2, 275–287, (1931).

[56] M.F. Spiegel and C.S. Watson, "Factors in the discrimination of tonal patterns. III. Frequency discrimination with components of well-learned patterns," *J. Acoust. Soc. Amer.* 69, 223–230, (1981).

[57] I.B. Thomas, P.B. Hill, F.S. Carroll, B. Garcia, "Temporal order in the perception of vowels," *J. Acoust. Soc. Amer.* 48, 1010–1013, (1970).

[58] R.M. Warren, *Auditory Perception; A New Synthesis*, NY: Pergamon, (1982).

[59] R.M. Warren, C.J. Obusek, R.M. Farmer, R.P. Warren, "Auditory sequence: Confusion of patterns other than speech and music," *Science* 164, 586–587, (1969).

[60] C.S. Watson, H.W. Wroton, W.J. Kelly, C.A. Benbassat, "Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals," *J. Acoust. Soc. Amer.* 57, 1175–1185, (1975).

[61] C.S. Watson, W.J. Kelly, H.W. Wroton, "Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty," *J. Acoust. Soc. Amer.* 68, 1176–1186, (1976).

[62] D.L. Wessel, "Timbre space as a musical control structure," *Computer Music J.* 3, 45–52, (1979).

[63] V.W. Zue, "Acoustic characteristics of stop consonants: a controlled study," Bloomington: Indiana Univ. Linguistics Club, (1980).

# PROCEEDINGS OF THE CONFERENCE

# ON THE PLANNING AND PRODUCTION OF SPEECH

# IN NORMAL AND HEARING-IMPAIRED INDIVIDUALS:

## A SEMINAR IN HONOR OF
## S. RICHARD SILVERMAN

*Edited by*

JUDITH L. LAUTER, PH.D.

with the editorial collaboration of:

Craig Bowe, Ph.D.
Don?ld R. Calvert, Ph.D.
Kenneth Grant, M.A.
Roanne Karzon, Ph.D.
James D. Miller, Ph.D.
Janet Weisenberger, Ph.D.

# PREFACE

Although oral education of hearing-impaired children and adults is based on a tradition that is centuries old, the science of speech production in hearing or hearing-impaired individuals is a young one. In the last few years interest in the details of speech planning and production has blossomed, drawing on new technologies such as electromyography, and new strategies such as analysis of speech errors.

The Silverman Seminar on the Planning and Production of Speech was organized in October 1983 to bring together two groups of experts: (a) those with experience in the day-to-day exigencies of teaching hearing-impaired children to speak, and (b) those whose research is directed to understanding the processes of speech planning and production in speakers with normal hearing. A combination of formal paper presentations, and working-group papers and discussions, was formulated to encourage the interaction of teachers and researchers.

The result is a collection of data, observations, questions, and answers that should be of interest to a wide range of readers. Those who are interested in how language is organized both as a system and within the human brain should find these proceedings of interest, as should teachers of the hearing-impaired, and other therapists who seek to understand the variety of disorders that can occur within the complicated interactive systems used to produce speech.

Judith L. Lauter, Ph.D.

# CONTENTS

Chapter 10

# RESPIRATORY FUNCTION IN SPEECH PRODUCTION BY NORMALLY-HEARING AND HEARING-IMPAIRED TALKERS: A REVIEW

JUDITH L. LAUTER

*Central Institute for the Deaf, St. Louis, MO*

Current models of speech production usually posit a fairly abstract semantic stage at one extreme, and a fairly specific stage of muscle-fiber control at the other, all referred to the sounds produced. In between, one or more intermediate stages are usually described, with various specifications. It might be suggested that one of these intermediate stages involves control of respiration, specifically, "breathing-for-speech." As we will see, control of respiration can be seen as involved intimately in a range of speech behaviors, from the articulation of phonemes to the rhythmic structuring of sentences. Also, problems that the severely hearing impaired have with speech production may be related to knowing how to breathe while talking.

## DYNAMICS OF SPEECH BREATHING

It has been known for some time (e.g., Stetson, 1951) that speech-breathing is somewhat different from quiet or "tidal" breathing. Borden and Harris (1980) note that more air is inspired during breathing for speech and the proportions of the cycle devoted to expiration and inspiration are very different (Figure 1). Hixon and colleagues (1973, 1976, 1982) have
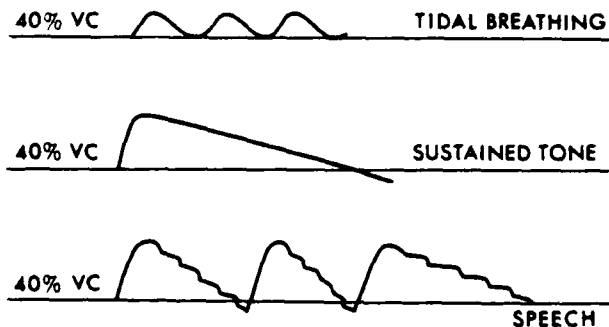


FIGURE 1. Use of lung capacity and rate of breathing compared for three different types of respiration: tidal (quiet) breathing, sustained vocalization, and normal speech. Lung capacity is indicated on the ordinate, and is shown in relation to 40% vital capacity, the lung volume at the end of a quiet expiration. Breathing rate is shown along the horizontal (time) axis, and is indicated by the relative slopes of inspiration versus expiration. (Reprinted with permission from Borden & Harris, 1980).

demonstrated a number of distinctions between quiet and speech-breathing.

These differences can be quite dramatic. Von Euler (1982) reports that while the muscles of the diaphragm continue to be active through about one half the expiration phase of quiet breathing, in speech-breathing, the muscles relax completely at the onset of expiration. Also, the metabolic reaction that occurs when subjects consciously hyperventilate without talking does not follow the hyperventilation that accompanies speech. Von Euler goes on to suggest that control of the two kinds of breathing may be partially separated in the CNS, quiet breathing depending on structures restricted to the brainstem and spinal cord, while voluntary breathing involves control centers in the cortex and basal ganglia as well (see also Abbs & Cole, 1982).

These and other data suggest that breathing-for-speech involves a set of motor skills that children must learn if they are to produce speech that sounds normal. Stetson (1951) presented his studies on speech production as an analysis of a set of "skilled movements." How might we think of the details of this skill as it relates to speech production? In a chapter published in 1973, Ron Netsell suggested that a useful description of the set of body structures used in speech production, the "speech apparatus," was as a system designed for generating and valving an airstream. The acts of control and coordination usually described with reference to the sounds thus produced, in this view are defined according to the effects on air flow and air pressure through the system.

Netsell went on to describe changes in airflow and pressure correlated with a range of linguistic events, from segmental to suprasegmental (prosody). He divided the "speech apparatus" into nine components (see Figure 2), and noted that the control of segmental aspects of speech in terms of this system required steady pressure maintained by the lower components, controlled modulation of the laryngeal "valve," modulation movements within the upper vocal tract, and extremely fine temporal-spatial coordination of all components—a coordination which must be able to comprehend within the same time frame the action of abdominal muscles as well as movements of the tip of the tongue.

Netsell noted how prosodic aspects of speech could be described in terms of the same components, with "valving" muscles and generated air pressure working together to achieve intonation—valve timing acting to achieve rhythm
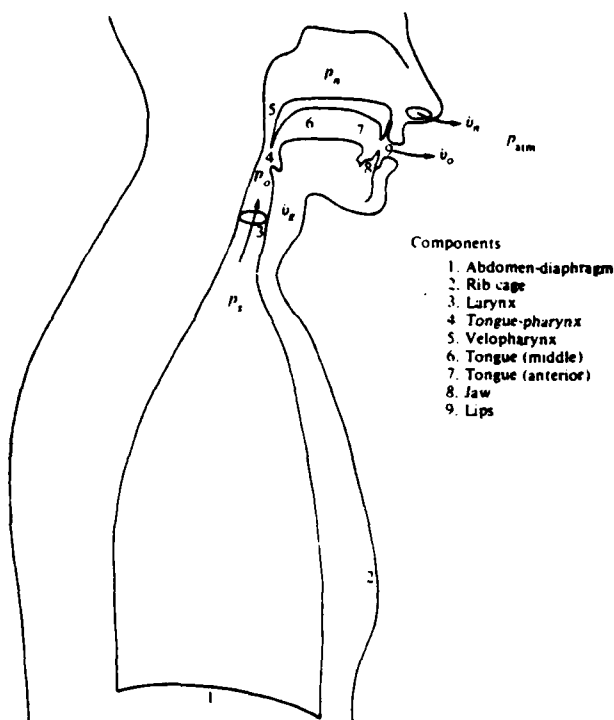
FIGURE 2. The nine components of the speech apparatus as described by Netsell (1973). The symbols /v/ and /p/ indicate various points where air volume and air pressure may be measured and compared. (Reprinted with permission from Netsell, 1973).

prasegmental aspects of speech—they read as though taken from Netsell's list of the speech aspects that depend intimately on temporal coordination of the nine speech-apparatus components. Deaf talkers' timing and rhythm are often abnormal in that these speakers may not provide clear duration distinctions between stressed and unstressed syllables, and segmental durations can be inaccurate. Also, pitch and intonation may be affected, in that base-line fundamental frequency is often too high, there is little variation in fundamental frequency, and modulations of intensity seem to be substituted for variations in pitch.

## Respiration and Speech Planning

It is possible that Netsell's approach to speech production may bear both on questions regarding the planning and production of speech, as well as the problems of individuals with handicaps such as motor disorders or hearing impairment. First, emphasis on this "intermediate," perhaps underlying skill of speech-breathing—defined in Netsell's broad terms of *airstream generation* and *modulation*—may serve as a guide for studying the physiology of speech acts. Interrelations between neural control centers and patterns of movement control may be suggested by this approach that would not emerge from thinking only about the sounds produced. For example, poor use of muscles of the torso for controlling subglottal pressure may have direct effects (perhaps via open-loop feed-forward connections) on control of the larynx, which could result in abnormalities in voice pitch. Von Euler (1982) has pointed out that the cerebellum may use the input it receives from both the larynx and the lungs to coordinate laryngeal and lower respiratory motor activities in phonation.

Certainly current concepts of motor control such as "heterarchical organization" (cf. Turvey, 1982) are compatible with this view of speech production. For example, details of motor action occurring against a background of general system "tuning" might be exemplified by the pulsed actions of the intercostal muscles timed against the background of other muscles acting to maintain subglottal pressure at a generally constant level throughout an utterance. MacNeilage's (Chapter 4) concept of "frame/content" organization could be illustrated by individual gestures of different valves programmed to match details of an utterance stress contour, and Bernstein's (1967) idea of "interactive coordinative structures" could be used to describe the interaction of the larynx and sublaryngeal structures to maintain subglottal pressure in the face of laryngeal actions such as opening and closing for segmental differentiation.

With developments in technology, it has become possible to perform noninvasive studies of the speech respiratory activity of normal and hearing-impaired individuals. Woldring (1968) used pneumographs (re: thorax and abdomen) to compare breathing patterns in one normal and two deaf children, from 10 to 12 years of age. He reported that during phonation the deaf subjects showed an absence of controlled expiration, with either insufficient ventilation or hyperventilation. He suggested that their poor control was due to the lack of auditory feedback; Woldring noted that "deaf glassblowers, in whom the feedback process is visual and not disturbed," show

and pitch control at both the segmental and phrase level, and subglottal pressure, extent of movements within the vocal cavity, and contact force of the different valves changing with the amount of effort of an utterance.

## Speech-Breathing and Deaf Speech

It is interesting to compare the linguistic effects of the actions of such a system with the characteristics of typical "deaf speech" such as described by Nickerson (1975) and Osberger and McGarr (1982). Nickerson divides these characteristics into classes that can be easily related to the segmental/suprasegmental distinction used by Netsell. Nickerson notes that deaf speakers often have poor articulation, including substandard velar control, a restricted range of F2 variation, problems with voiced/voiceless distinctions and with "continuous phonation." Many of these characteristic difficulties in deaf speech could be described as inadequate control (perhaps in terms of poorly learned control constraints) of Netsell's "valves"—the velum, the tongue, the vocal folds.

The same could be said for Nickerson's list of the characteristics of deaf "voice quality"—nasality (velar valve), breathiness (laryngeal valve), inappropriate loudness (perhaps compensation using changes in intensity controlled at the larynx instead of changes in fundamental frequency managed there), and durational distortions. Even more suggestive are Nickerson's deaf-speech characteristics having to do with su-

good control of respiration skills needed in glassblowing. Forner and Hixon (1977), using the kinematic procedure developed by Hixon, et al. (1973), reported a study of 10 young male deaf students. Two pairs of magnetometer coils were used to measure movements of the chest and abdomen during a variety of respiratory maneuvers, including quiet breathing, and breathing during a series of speech tasks. The authors concluded that although the deaf speakers showed quiet breathing patterns that were within normal limits, their speech breathing was generally deviant. Departure from normal behaviors included: fewer syllables per breath than normals, less air inspired with each breath than normals, higher volume of air per syllable than normals, and inspirations taken at linguistically irrelevant points. Whitehead (1983) used similar measurement techniques to study 15 young deaf males, whose speech was rated as semiintelligible or unintelligible. He reported results similar to those seen by Forner and Hixon (1977), and suggested that speech intelligibility might be affected by a speaker's respiratory skill. Specifically, such practices as initiating speech at low lung volumes, and continuing speech beyond the lower limit of tidal breathing, could contribute directly to listeners' difficulty with comprehension.

*Aerodynamic Feedback for Deaf Talkers*

Second, as Woldring (1968) suggested, it is possible that individuals who cannot hear the effect of actions of the "speech apparatus" may benefit from feedback directly related to its aerodynamics. Forner and Hixon (1977) included in their report a final study where they showed one of the deaf talkers the display that formerly only the experimenters had seen; and taught the subject how movements of his torso could affect the tracing. After a few minutes of working with the display, the hearing-impaired speaker learned to: (a) produce a speech-breathing pattern more like that of a normal speaker, and (b) as a side effect, without direct attention by experimenters or subject, lower his abnormally high "deaf voice pitch" to a normal level.

Certainly the importance of "breathing exercises" is cited in the oldest treatises on oral education of the deaf. However, examination of these descriptions reveals a lack of understanding about the intimate relations between gestures within the respiratory tract and segmental *and* suprasegmental characteristics of speech. It is possible that the general failure in teaching the deaf to produce normal speech is based in part on the failure to teach them the motor skills involved in breathing for speech. Forner and Hixon (1977) reported that some of their hearing-impaired subjects said that they were only "taught to make speech sounds, never to breathe in a different way for speech than for quiet breathing." As we have suggested, the range of relevant motor skills involve a variety of details, from knowing that more air needs to be inspired for speech than for tidal breathing to being aware of the effect of leaving the velopharyngeal valve open.

## CONCLUSION

In the past, conclusions about planning for speech production have been drawn from observations of speech dysfunction, as in aphasia, dysarthria, and spontaneous speech errors (for the latter cf. Fromkin, 1973; Shattuck-Hufnagel, 1985). It is possible that observations of the aerodynamics of deaf speech, both before and after relevant instruction, may provide new evidence for the stages of planning, centers of control, and details of coordination that are involved in creating the disturbances in the air that we hear as speech.

## REFERENCES

ABBS, J. H., & COLE, K. J. (1982). Consideration of bulbar and suprabulbar afferent influences upon speech motor coordination and programming. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson (Eds.), *Speech motor control* (pp. 75–93). Oxford: Pergamon Press.

BERNSTEIN, N. (1967). *The coordination and regulation of movements.* Oxford: Pergamon Press.

BORDEN, G. J., & HARRIS, K. S. (1980). *Speech science primer.* Baltimore: Williams & Wilkins.

FORNER, L. L., & HIXON, T. J. (1977). Respiratory kinematics in profoundly hearing-impaired speakers. *Journal of Speech and Hearing Research, 20,* 373–408.

FROMKIN, V. A. (1973). *Speech errors as linguistic evidence.* The Hague: Mouton.

VON EULER, C. (1982). Some aspects of speech breathing physiology. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson (Eds.), *Speech motor control* (pp. 95–103). Oxford: Pergamon Press.

HIXON, T. J. (1982). Speech breathing kinematics and mechanism inferences therefrom. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson (Eds.), *Speech motor control* (pp. 75–93). Oxford: Pergamon Press.

HIXON, T. J., GOLDMAN, M., & MEAD, J. (1973). Kinematics of the chest wall during speech production: Volume displacements of the rib cage, abdomen, and lung. *Journal of Speech and Hearing Research, 16,* 78–115.

HIXON, T. J., MEAD, J., & GOLDMAN, M. D. (1976). Dynamics of the chest wall during speech production: Function of the thorax, rib cage, diaphragm, and abdomen. *Journal of Speech and Hearing Research, 19,* 297–356.

NETSELL, R. (1973). Speech physiology. In F. D. Minifie, T. J. Hixon, & F. Williams (Eds.), *Normal aspects of speech, hearing, and language* (pp. 211–234). Englewood Cliffs, NJ: Prentice-Hall.

NICKERSON, R. S. (1975). Characteristics of the speech of deaf persons. *Volta Review, 77,* 342–362.

OSBERGER, M. J., & McGARR, N. S. (1982). Speech production characteristics of the hearing impaired. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice, Vol. 8* (pp. 221–283). New York: Academic Press.

STETSON, R. H. (1951). *Motor phonetics* (2nd ed.). Amsterdam: North-Holland Publishing.

TURVEY, M. T. (1982). Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: toward an ecological psychology,* (pp. 211–265). Hillsdale, NJ: Lawrence Erlbaum Associates.

WHITEHEAD, R. L. (1983). Some respiratory and aerodynamic patterns in the speech of the hearing impaired. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired* (pp. 97–116). Baltimore: University Park Press.

WOLDRING, S. (1968). Breathing patterns during speech in deaf children. *Annals of the New York Academy of Sciences, 155,* 206–207.

# Tonotopic organization in human auditory cortex revealed by positron emission tomography *

Judith L. Lauter [1,**], Peter Herscovitch [2], Craig Formby [3] and Marcus E. Raichle [2]

*[1] Central Institute for the Deaf, St. Louis, MO 63110, [2] Edward Mallinckrodt Institute of Radiology, McDonnell Center for Studies of Higher Brain Function, Department of Neurology and Neurological Surgery, Washington University School of Medicine, St. Louis, MO 63110, and [3] Departments of Communicative Disorders and Neurology, J.H. Miller Health Center, University of Florida, Gainesville, FL 32610, U.S.A.*

Positron emission tomography (PET) was used to map alterations in local neuronal activity induced in human primary auditory cortex by pure-tone stimulation. Patterns of blood flow were observed in specific regions on the superior temporal plane showing systematic changes in activity depending on the frequency of a stimulating pure tone. The orientation of these regions agrees well with data for non-human primates.

tonotopic organization, human primary auditory cortex, positron emission tomography, regional cerebral blood flow

## Introduction

It has been known for some time that primary auditory cortex in a number of animals exhibits tonotopic organization. The earliest reports were by Licklider and Kryter (1942) and Walzl and Woolsey (1943). Subsequent descriptions have been published by a number of authors. However, the invasive nature of most physiological techniques has prevented researchers from determining whether human brains exhibit a similar organization.

Recently two techniques have become available that provide such a capability. The first of these, magnetoencephalography, accomplished by means of a 'superconducting quantum interference device' (SQUID: cf. Kaufman and Williamson, 1980), has been applied to the study of responses in

human brains to pure-tone stimulation (Elberling et al., 1982; Romani et al., 1982). These measurements were made with single-channel devices, and were thus limited to monitoring brain responses along a single dimension: depth beneath the skull. The second technique is positron emission tomography (PET). PET provides quantitative measurements of regional cerebral blood flow (rCBF) and metabolic rate for oxygen and glucose in the human (Raich.e, 1983), comprising a three-dimensional representation of the brain, with resolution better than 2 cm in both the horizontal and vertical dimensions. Under normal circumstances PET measurements are thought to reflect the local rate of neuronal activity (Yarowsky and Ingvar, 1981). The relationship between neuronal function, metabolic rate, and blood flow underlies the extensive use of positron and single-photon emission imaging for functional-anatomical mapping of the brain (e.g., Fox and Raichle, 1984; Phelps et al., 1981a; Reivich, 1982; Roland et al., 1980; Roland, 1982).

Published PET studies of the auditory system of humans have for the most part involved rather imprecise auditory stimulation, e.g., music

(Carmon et al., 1975) and stories (Phelps et al., 1981b). Although positive responses to such stimuli have been observed with PET, the stimulus, subject, and presentation variables responsible for the responses observed are unclear. In this study we examined the responses of human auditory cortex to pure-tone stimulation at two frequencies, using PET to measure changes in local CBF. We used a sound delivery system that ensured isolated stimulation of the two ears (see below), a stimulus protocol that involved repetitive measurements in the same subject (Fox and Raichle, 1984), and an anatomical localization scheme that is free of observer bias (Fox et al., 1984).

### Materials and Methods

#### PET

Positron emission tomography was performed with a PETT VI system (Ter-Pogossian et al., 1982; Yamamoto et al., 1982). Data were recorded simultaneously for 7 slices with a center-to-center separation of 14.4 mm. All studies were done in the low-resolution mode, giving an in-plane (i.e., transverse) reconstructed resolution of about 12.4 mm in the center of the field of view and a slice (axial) thickness of 13.9 mm at the center.

Each scan was 40 s in length, and was performed following the intravenous bolus injection of about 10 ml of saline containing 55–80 mCi of $^{15}$O-labelled water (half life: 123 s). CBF (ml/(min × 100 g)) was calculated using a PET adaptation of the Kety tissue autoradiographic technique previously described and validated in our laboratory (Herscovitch et al., 1983; Raichle et al., 1983).

#### Stimuli

The circuit for generating and presenting the pure tones included a General Radio 1310A oscillator, an electronic switch and pulse generator (built at Central Institute for the Deaf), and a Hewlett-Packard 350D attenuator. Tones were monitored using a Monsanto 113A counter, a Telequipment S54A oscilloscope and a Hewlett-Packard 400GL voltmeter. Sounds were presented to the subject through Knowles ED-1912 insert receivers set in plastic tubing connected to shaped ends that fit into the snap rings of a set of standard earmolds. The frequency response of this



Sound delivery system frequency response (---) compared with transfer function of outer ear (——)
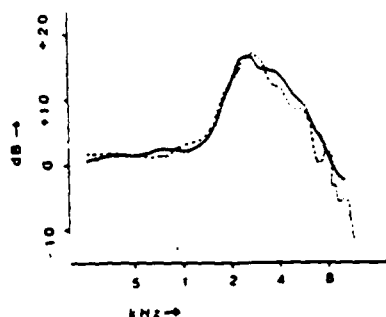
Fig. 1. Frequency response of the sound delivery system compared with the filter characteristics of the outer ear (Shaw, 1974). This configuration ensures that the sound presented to the eardrum is essentially the same as if the sound were presented in the field.

system was modified * to mimic the filter characteristics of the pinna and outer-ear canal (Shaw, 1974). The result of these modifications is shown in Fig. 1. This configuration was designed to ensure that sounds would be presented in an 'ecologically valid' form to the eardrum. It is used for all auditory stimulation on the PETT VI.

Tones of 500 Hz and 4 kHz were used for testing. Tones were pulsed with a duty cycle of 50%, approximately 500 ms on/off, with a rise/fall time of 50 ms. The subject's threshold for each frequency tested was determined just prior to scanning for that frequency. All tones were presented at 50 dB SL, monaurally to the right ear. Four subjects were tested with the 500 Hz tone presented during the first pair of experimental scans, and the 4 kHz tone presented during the second pair. One subject heard the tones in the reverse order, and one of the first four subjects was tested in an additional session, comprising an

---

* The sound delivery system was designed and built at Central Institute for the Deaf by Arnold Heidbreder. It combines a parametric equalizer, Knowles ED-1912 insert receiver, 2 mm of receiver tubing, 35 mm of No. 13 tubing, a 1500 Ω acoustic resistor, a right-angle plastic snap, and a standard stock earmold, to mimic the frequency response of the outer ear as reported by Shaw (1974). Further details of the system are included in Central Institute for the Deaf Periodic Progress Report No. 25, pp. 31–32.

initial condition where the two frequencies alternated, followed by a second condition with the 4-kHz tone presented alone.

## Subjects

Subjects were five normal young volunteers and each was paid $50 for each half-day session. Prior to testing, each subject received an orientation visit to the laboratory, where the devices were shown, all procedures explained, and a consent form was read and signed.

Subject preparation preceding each session included the percutaneous insertion of a radial arterial catheter, under local anesthesia, to permit frequent sampling of arterial blood, and the insertion of an intravenous catheter in the opposite arm for isotope injection. The head was positioned with a special head holder which utilized an individually molded plastic face mask to prevent movement during the study. A laser permanently attached to the wall projected a line onto the mask that corresponded to the position of the lowest PET slice. A lateral skull radiograph with this line marked by a radiopaque wire provided a record of the subject's exact position in relation to the PET slices. The overlapping position of radiopaque markers placed in the external auditory canals (the earmold rings) confirmed that the head was not rotated about the anterior–posterior or vertical axes. After the head was in place, a transmission scan used for individual attenuation correction was performed with a ring source of activity containing germanium-68/gallium-68. During each PET scan the room was darkened and the subject's eyes were covered with gauze pads. Ambient noise during each scan was limited to the sound of cooling fans for the electronic equipment.

Following subject preparation, the protocol for pure-tone testing was begun. In order to obtain a 14-slice representation of the brain, each condition was tested with two head positions: one at zero position within PETT VI, and a second at 7.2 mm rostral to the original position. An eight-scan session consisted of an initial pair of control scans (no auditory stimulation other than ambient sound), a pair of experimental scans (tone of $x$ frequency), a second pair of experimental scans (tone of $y$ frequency), a final pair of control scans. For each experimental scan, the sound was turned

on approximately 1 min prior to isotope injection, and was presented throughout the scan; thus total presentation time was approximately 2 min.

## Anatomical localization

In order to determine where in the three-dimensional complex of data to look for responses of primary auditory cortex, we used an anatomical localization scheme developed in our laboratory (Fox et al., 1984) that is independent of the appearance of the CBF images. This method yields both slice number and coordinates in the transverse plane for a predicted region of interest (ROI) selected from a standard stereotaxic atlas of the human brain (Talairach et al., 1967).

The transverse gyri of the superior temporal plane which comprise human primary auditory cortex may consist of from one to five gyri per side, covering an average area on each superior temporal plane of about 1200 mm$^2$ (Campain and Minckler, 1976; Celesia, 1976; Celesia and Puletti, 1969; Galaburda and Sanides, 1980). Using the stereotaxic atlas coordinates for the center of this region (vertical: 13 mm above the frontal–orbital line; right–left: 5.6 cm lateral to brain midline; anterior/posterior: 1.3 cm posterior to the AP midpoint), and our anatomical localization procedure cited above, we identified the center of the target region on the appropriate PET slice for each subject. The entire representation of the AI region then was designated as a 4 × 4 cm square on this slice.

Tomographic images from each subject were then to used to create 'percent-difference images' (Fox and Raichle, 1984), comparing control and experimental conditions. These images are based on blood-flow values normalized to control for global changes in blood flow occurring between scans, and to highlight areas of maximum change from control to stimulated condition that occur independent of any global changes in CBF.

## Results

Examination of activity changes within the estimated region of primary auditory cortex for each hemisphere of each subject revealed systematic shifts of the area of maximum change from condition to condition. In each subject, max-

imum change always occurred in the left-hemisphere AI region (i.e., contralateral to stimulation). Also, for each subject, the contralateral region of greatest activity change during stimulation with the 500-Hz tone was more lateral and anterior, and the region that responded best (with one exception: see below) to the 4-kHz tone was more medial and posterior. For hemispheric comparisons, right-hemisphere mirror images of the left-side regions of change were selected and analyzed. This method was taken as a conservative first step to representing the pattern of changes observed; analysis by eye indicated that there was no consistent pattern of ipsilateral response comparable to that seen contralaterally. Quantitative measures of the changes in each region were obtained by placing a 1.5 cm-square cursor over each area of maximal change. Values of change in each of the
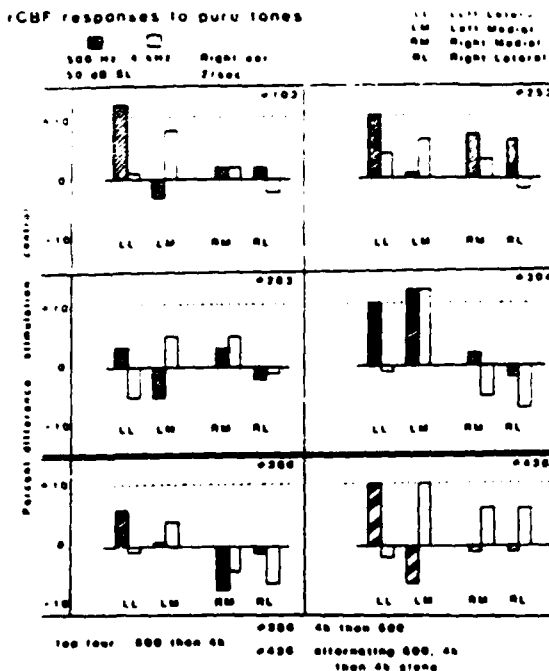


Fig. 2. Percent change in blood flow seen in four primary auditory cortex regions (left lateral, left medial, right medial and right lateral) in response to pure tone stimulation on each of six test sessions. The top four sessions involved testing a 500 Hz tone before a 4 kHz tone. In session No. 386 the testing order of the frequencies was reversed, and in session No. 436 a condition of alternating frequencies was followed by the 4 kHz tone alone. The subject tested in sessions Nos. 283 and 436 was the same individual.
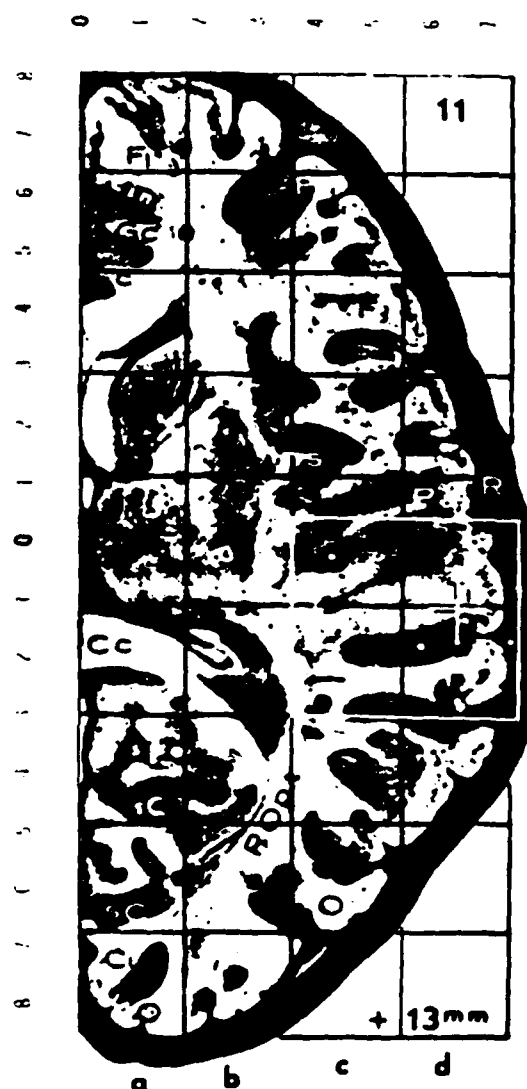
four regions identified in each subject (i.e., two on the left, and their mirrors on the right) are shown in Fig. 2.

The top four subjects in the figure were tested with the protocol where the 500-Hz tone was presented during the first two experimental scans, and the 4-kHz tone during the second two scans. Clearly there are individual differences in these responses, but the general pattern is the same: (1) there is more contralateral response than ipsilateral; (2) on the left, the more lateral regions identified for analysis respond better to the low tone than to the high tone; and (3) the more medial contralateral regions responded better to the high tone than to the low. The exception is the medial region identified for Subject 304 which responded equally well to both frequencies.

In the two sessions shown at the bottom of the figure, subjects were tested on somewhat different protocols. Subject 386 heard the high tone before the low; the pattern of response is much like that seen in the other four subjects. Subject 436 heard first a condition in which the two tones alternated (each preset to be 50 dB SL), and a second condition of 4 kHz alone. Note that for this subject, the medial region responds well to the high tone (as in all other subjects), but during the alternating condition, there was best response in the lateral region. This suggests that during the alternating condition, rCBF response was dominated by the low tone. It should also be noted that the subject of sessions 283 and 436 was the same individual. The two sessions were run 10 months apart.

For the six sessions, a mixed-design analysis of variance (one Between factor of region, one Within factor of frequency) was used to compare percent change of rCBF in the left and right ROIs under the two frequency conditions. For the left-side (contralateral) ROIs, the interaction between frequency and region was significant at the 0.01 level ($F = 44.1$; d.f. $= 1, 10$; significant $F$ at $P < 0.01 = 10.00$). There were no significant differences on the left due to either region alone ($F = 0.0019$; d.f. $= 1, 10$; signif. $F$ at $P < 0.05 = 4.96$), or tone frequency alone ($F = 0.41$; d.f. $= 1, 10$). For the right-side (ipsilateral) regions, there were no significant differences in CBF percent change with regard to region or frequency or interactions ($F = 0.85$ for region, 0.18 for frequency,

Fig. 3. Regions of maximum change in rCBF, reconstructed in the horizontal plane on a standard brain (Talairach et al., 1967). Left–right values have been transformed for use on this right-hemisphere slice. Areas of responses with indicated mean and S.D. over six sessions responding best to the low tone are represented by the lateral cluster of points: areas responding best to the high tone are represented by the more medial points.

0.77 for the interaction; d.f. = 1, 10; signif. $F$ at $P < 0.05$ = 4.96). Mean values of left- and right-side CBF changes are shown in Table I.

The locations in the transverse plane of regions

**TABLE I**

MEAN CHANGES IN rCBF (EXPRESSED AS PERCENT CHANGE re CONTROL) FOR TWO REGIONS IN THE LEFT HEMISPHERE AND TWO IN THE RIGHT. CHOSEN INDIVIDUALLY FOR 5 SUBJECTS IN EACH OF 6 SESSIONS

Monaural stimulation to the right ear (details in text).

| Region | Tone in Hz | Left-side regions ($\bar{x}$ ± S.D.) | Right-side regions ($\bar{x}$ ± S.D.) |
|--------|-----------|------------------|-------------------|
| Lateral | 500 | +8.67 ± 3.01 | +1.00 ± 5.48 |
|  | 4000 | −0.83 ± 3.31 | +1.50 ± 4.89 |
| Medial | 500 | −0.17 ± 6.74 | +0.10 ± 3.85 |
|  | 4000 | +7.67 ± 3.01 | −1.83 ± 4.40 |

responding to the two tones, with the mean and standard deviation over the six sessions, are given in Table II. These areas, with the left–right values transformed for representation on the right-hemisphere atlas (Talairach et al., 1967) slice are shown in Fig. 3. The cluster of points more anterior and lateral represent regions responding better to the 500 Hz pure tone (mean and 1 S.D. in both AP and left–right dimensions are indicated), and the more medial points represent regions responding better to the 4 kHz tone.

**Discussion**

This study clearly demonstrates that human auditory cortex responds to pure-tone stimulation in a tonotopic manner. The orientation of the

**TABLE II**

LOCATIONS IN THE TRANSVERSE PLANE OF REGIONS RESPONDING TO PURE TONES OF 500 Hz AND 4 kHz: 6 SESSIONS (5 INDIVIDUALS)

Values are given in centimeters relative to the anterior/posterior midpoint and the brain midline as shown in the atlas of Talairach et al. (1967), p. 131 (cf. Fig. 3).

|  | cm posterior to AP midpoint | | | cm lateral to brain midline | | |
|--|--|--|--|--|--|--|
| 500 Hz | 0.0 | 1.0 | 1.6 | 5.9 | 6.7 | 6.8 |
|  | 0.5 | 1.3 | 1.7 | 6.7 | 6.8 | 7.3 |
| $\bar{x}$ ± S.D. | | 1.0 ± 1.3 | | | 6.7 ± 0.9 | |
| 4 kHz | 0.3 | 1.1 | 1.7 | 3.7 | 3.7 | 4.3 |
|  | 0.9 | 1.1 | 2.0 | 3.7 | 4.1 | 4.3 |
| $\bar{x}$ ± S.D. | | 1.2 ± 1.2 | | | 4.0 ± 0.6 | |

observed responses agrees well with data based on electrophysiological measurements in other primates (e.g., Merzenich and Brugge, 1973).

Since the early 1940's (cf. Walzl and Woolsey, 1943), it has been known that primary auditory cortex in animals other than humans shows tonotopic organization. The relative orientation of cells responding best to low as compared to higher frequencies remains fairly constant in the face of evolutionary changes in orientation of the temporal lobe.

Because of the technology involved, observations of neural activity in restricted cortical areas have been limited to non-humans. However, developments in both autoradiography and neuromagnetic techniques have made it possible to observe changes in human brains in response to stimulation. There have been attempts (Elberling et al., 1982; Romani et al., 1982) to use magnetoencephalography to study responses to pure-tone stimulation in human primary auditory cortex. In both cases, the experimenters were able to discern different responses to lower and higher tones, with responses aligned along a single dimension: depth below the skull. Since the hypothesized orientation of Heschl's gyrus is from superficial to deep, this happens to be an appropriate dimension for studying tonotopic mapping in humans.

A number of previous PET studies have attempted to measure the responses in human brains to auditory stimulation. The majority of experiments reported have made use of sounds more complex than pure tones. A sample of sounds used includes: tone sequences (e.g., Phelps et al., 1981b; Roland et al., 1981), noises (e.g., Miyazaki, 1971, 1978; Hiroshige and Iwahara, 1978; Lassen et al., 1978; Knopman et al., 1980; Lassen and Larsen, 1980), music (Carmon et al., 1975), isolated words (Carmon et al., 1975; Maximilian, 1980; Lassen et al., 1978; Knopman et al., 1980) and narrative stories (Carmon et al., 1975; Larsen et al., 1977; Reivich et al., 1979; Alavi et al., 1981; Greenberg et al., 1981; Phelps et al., 1981b; Maximilian, 1980). In all of these experiments different proportions of response were seen in temporal, parietal, and/or frontal cortex, depending on the task given the subjects. It is to be expected that responses to the spectrally and temporally complex sounds tested in these studies should be much more dif-

fuse than responses to pure tones. Thus stimulus design of these experiments precluded observation of frequency-specific responses. In addition, limited resolution and the difficulty of precisely locating anatomical structures corresponding to details of the physiological PET images restricted the experimenters to poorly defined descriptions of activated regions of cortex.

Complex perceptual tests such as those mentioned above comprise numerous variables whose impact upon CBF and brain metabolism is unknown. Some of these are related to the stimulus (which and how many dimensions are manipulated; the complexity of the stimulus; the complexity of stimulus presentation, e.g., monaural vs. dichotic) and others are related to the subject (whether the subject passively receives the stimuli; whether a task is assigned; the complexity of the task). Systematic work on the effect on CBF and brain metabolism of stimulus and subject variables must be pursued before more complex activation paradigms can be properly designed, executed or interpreted.

**References**

Alavi, A., Reivich, M., Greenberg, J., Hand, P., Rosenquist, A., Rintelmann, W., Christman, D., Fowler, J., Goldman, A., MacGregor, B. and Wolf, A. (1981): Mapping of functional activity in brain with 18-F-fluoro-deoxyglucose. Seminars Nuclear Med. 11, 24–31.

Campain, R. and Minckler, J. (1976): A note on the gross configurations of the human auditory cortex. Brain Lang. 3, 318–323.

Carmon, A., Lavy, S., Gordon, H. and Portnoy, Z. (1975): Hemispheric differences in rCBF during verbal and non-verbal tasks. In: Brain Work, Alfred Benzon Symposium VIII, pp. 414–423, Munksgaard, Copenhagen.

Celesia, G.G. (1976): Organization of auditory cortical areas in man. Brain 99, 403–414.

Celesia, G.G. and Puletti, F. (1969): Auditory cortical areas of man. Neurology 19, 211–220.

Elberling, C., Bak, C., Kofoed, B., Lebech, J. and Saermark, K. (1982): Auditory magnetic fields: source location and 'tonotopical organization' in the right hemisphere of the human brain. Scand. Audiol. 11, 61–65.

Fox. P.T. and Raichle. M.E. (1984): Stimulus rate dependence of regional cerebral blood flow in human striate cortex, demonstrated by positron emission tomography. J. Neurophysiol. 51. 1109-1120.

Fox. P.T.. Perlmutter. J.S and Raichle. M.E. (1984). A stereotactic method of anatomical localization for positron emission tomography. J. Computer Assisted Tomogr. 9. 141-153.

Galaburda. A.M. and Sanides. F. (1980) Cytoarchitectonic organization of the human auditory cortex. J. Comp. Neurol. 190. 597-610.

Greenberg. J.H.. Reivich. M.. Alavi. A.. Hand. P.. Rosenquist. A.. Rintelmann. W.. Stein. A.. Tusa. R.. Dann. R.. Christman. D.. Fowler. J.. MacGregor. B. and Wolf. A (1981): Metabolic mapping of functional activity in human subjects with the 18-F-fluorodeoxyglucose technique. Science 212. 678-680.

Herscovitch. P.. Markham. J. and Raichle. M.E. (1983): Brain blood flow measured with intravenous H2-015. 1. Theory and error analysis. J. Nuclear Med 24. 782-789.

Hiroshige. Y. and Iwahara. S. (1978) Digital and cephalic vasomotor orientating responses to indifferent. signal. and verbal stimuli. Psychophysiology 15. 226-232.

Kaufman. L and Williamson. S.J (1980) The evoked magnetic field of the human brain. Ann. N.Y. Acad. Sci. 340. 45-65.

Knopman. D.S.. Rubens. A.B.. Klassen. A.C.. Meyer. M.W and Niccum. N (1980): Regional cerebral blood flow patterns during verbal and nonverbal auditory activation. Brain Lang. 9. 93-112.

Larsen. B.. Skinhoj. E.. Soh. K. Endo. H and Lassen. N A (1977). The pattern of cortical activity provoked by listening and speech revealed by rCBF measurements. Acta Neurol. Scand. Suppl 64. 268-269.

Lassen. N A. and Larsen. B. (1980) Cortical activity in the left and right hemispheres during language-related brain functions. Phonetica 37. 27-37.

Lassen. N.A. Ingvar. D.H. and Skinhoj. E. (1978): Brain function and blood flow. Sci. Am. 62-71.

Licklider. J.C.R. and Kryter. K.D (1942). Frequency-localization in the auditory cortex of the monkey. Fed. Proc 1. 51.

Maximilian. V.A (1980). Functional changes in the cortex during mental activation. Ph.D Dissertation. Laboratory of Neuropsychology. Dept. of Psychiatry. University Hospital. Lund. Sweden.

Merzenich. M.M. and Brugge. J.F. (1973): Representation of the cochlear partition on the superior temporal plane of the macaque monkey. Brain Res. 50. 275-296.

Miyazaki. M. (1971): Effect of undesirable sound (noise) on cerebral circulation. Jpn. Circ. J. 35. 931-936.

Miyazaki. M. (1978): Circulatory effect of jet noise. with special reference to cerebral circulation. Jpn. Circ. J. 42. 1019-1024.

Phelps. M.E.. Kuhl. D.E. and Mazziotta. J.C. (1981a): Metabolic mapping of the brain's response to visual stimulation: studies in humans. Science 211. 1445-1448.

Phelps. M.E.. Mazziotta. J.C.. Engel. J. Jr. and Kuhl. D.E.

(1981b): Metabolic response of the brain to visual and auditory stimulation and deprivation. J. Cereb. Blood Flow Metab. 1 (Supp 1). S467-468.

Raichle. M.E. (1983): Positron emission tomography. Annu. Rev. Neurosci. 6. 249-267.

Raichle. M.E.. Grubb. R.L.. Gado. M.H.. Eichling. J.O. and Ter-Pogossian. M.M. (1976): Correlation between regional cerebral blood flow and oxidative metabolism. Arch. Neurol. 33. 523-526.

Raichle. M.E.. Martin. W.R.W.. Herscovitch. P.. Mintun. M.A. and Markham. J. (1983) Brain blood flow measured with intravenous H2-015. II. Implementation and validation. J. Nuclear Med. 24. 790-798.

Reivich. M. (1982): The use of cerebral blood flow and metabolic studies in cerebral localization. In: New Perspectives in Cerebral Localization. pp. 115-143. Editors: R.A. Thompson and J.R. Green. Raven Press. New York.

Reivich. M.. Greenberg. J.. Alavi. A.. Christman. D.. Fowler. J. Hand. P.. Rosenquist. A.. Rintelmann. W. and Wolf. A (1979) The use of the 18-F-fluorodeoxyglucose method for mapping of functional neural pathways in man. Acta Neurol. Scand. Suppl. 72. 198-199.

Roland. P.E. (1982) Cortical regulation of selective attention in man: a regional cerebral blood flow study. J. Neurophysiol. 48. 1059-1078.

Roland. P.E.. Larsen. B. and Lassen. N.A (1980): Supplementary motor area and other cortical areas in organization of voluntary movements in man. J. Neurophysiol 43. 118-136.

Roland. P.E. Skinhoj. E. and Lassen. N.A. (1981) Focal activations of human cerebral cortex during auditory discrimination. J. Neurophysiol 45. 1139-1151.

Romani. G.L. Williamson. S.J. and Kaufman. L. (1982) Tonotopic organization of the human auditory cortex. Science 216. 1339-1340.

Shaw. E.A.G. (1974): The external ear. In: Handbook of Sensory Physiology. Vol V/1. Auditory System: Anatomy and Physiology (Ear). pp 455-490. Editors W.D. Keidel and W.D. Neff. Springer-Verlag. Berlin.

Talairach. J.. Szikla. G and Tournous. P (1967) Atlas d'Anatomie stereotaxique du Telencephale. Masson. Paris.

Ter-Pogossian. M.M.. Ficke. D.C.. Hood. J.T. Sr.. Yamamoto. M and Mullani. N.A (1982): PETT VI: Positron emission tomography utilizing cesium fluoride scintillation detectors. J. Computer Assisted Tomogr. 6. 125-133.

Walzl. E.M. and Woolsey. C.N. (1943) Cortical auditory areas of the monkey as determined by electrical stimulation of nerve fibers in osseus spiral lamina and by click stimulation. Fed. Proc. 2. 52.

Yamamoto. M.. Ficke. D.C and Ter-Pogossian. M.M. (1982): Performance study of PETT VI. a positron computed tomograph with 288 cesium fluoride detectors. IEEE Trans Nuclear Sci. NS29. 529-533.

Yarowsky. G.J. and Ingvar. D.H. (1981): Neuronal activity and energy metabolism. Fed. Proc. 40. 2353-2363.

# INDIVIDUAL DIFFERENCES IN AUDITORY ELECTRIC RESPONSES: COMPARISONS OF BETWEEN-SUBJECT AND WITHIN-SUBJECT VARIABILITY

## I. *Absolute latencies of brainstem vertex-positive peaks*

Judith L. Lauter and Robert L. Loomis

*From the Central Institute for the Deaf, Missouri, St. Louis, Missouri and the Dept. of Otolaryngology, Washington University School of Medicine, Missouri, St. Louis, Missouri, USA*

## ABSTRACT

*Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability. I. Absolute latencies of brainstem vertex-positive peaks. Lauter, J. L. and Loomis, R. L. (Central Institute for the Deaf, 818 S. Euclid and Dept. of Otolaryngology, Washington University School of Medicine, 660 S. Euclid, St. Louis, MO, USA.).*

*Scand Audiol 1986, 15 (167–172).*

Seven subjects were tested, each on eight separate sessions, for brainstem auditory evoked responses to monaural right, monaural left, and binaural stimulus presentations. Comparisons of between-subject vs. within-subject variability of the absolute latencies of vertex-positive peaks expressed in terms of the coefficient of variation indicate that: 1) within-subject stability is greater than between-subject stability for the five brainstem peaks; 2) between-subject variability shows significant differences due to peak but not to ear of presentation; 3) within-subject variability shows significant differences due to both peak and ear; 4) comparisons of within-subject variability over time show significant differences due to peak but not to time; 5) patterns of individual variation within the brainstem series are characterized by increases in stability of peak latencies over time, and by replicability of stability profiles over time. Other measures of latency and amplitude based on this series of responses are planned for a subsequent report.

## INTRODUCTION

Over the last few years, a number of authors have reported data regarding the range of normal variability in auditory evoked responses (AERs). For example, variability within groups of subjects has been studied by Thornton (1975), Rosenhamer et al. (1978), Kendall & Lawes (1978), Chiappa et al. (1979), Kjaer (1979), Stockard et al. (1979), Spreng (1980), Bergholtz (1981), Owen & Matsusaka (1982), and Rosenhamer & Holmkvist (1982). AER parameters have also been compared as a function of specific subject characteristics such as age or sex

(Beagley & Sheldrake, 1978; Goodin et al., 1978; Rowe, 1978; Stockard et al., 1979; Kjaer, 1980; Allison et al., 1983; Stockard et al., 1983) and as a function of mode of stimulus presentation: monaural versus binaural or left versus right (Blegvad, 1975; Ainslie & Boston, 1980; Dobie & Norton, 1980; Levine & McGaffigan, 1983).

This literature is concerned primarily with the normal range of *between-subject* variability in AERs. Currently lacking are comparable data regarding the range of *within-subject* AER variability. Although single-session test–retest measurements have been reported by Aunon & Cantor (1977), and Owen & Matsusaka (1982) [for cortical responses], and by Chiappa et al. (1979) and Edwards et al. (1982) [brainstem responses], no results of extensive repeated-measure testing have been reported. The present study was designed to provide data that would allow comparisons of between-subject vs. within-subject variability in the same group of subjects, for the first five vertex-positive peaks of the brainstem response.

## METHODS

Auditory brainstem responses (ABRs) were recorded from 7 normal young adult subjects, 4 females and 3 males. Each subject was tested on eight separate weekly sessions. All sessions for each subject were scheduled for the same time of day or the same day of the week. Each subject was screened for normal hearing threshold preceding each test session; on one occasion, a subject's threshold was elevated due to a mild middle-ear infection, and the session was rescheduled for the following week.

Each session included monaural right, monaural left, and binaural stimulation. Nine-millimeter silver disk recording electrodes were placed at Cz, A1, and A2. A ground was placed at Fpz. For monaural presentations, Cz was referenced to the ipsilateral ear; for binaural, Cz was referenced to linked earlobes. Stimuli were 100 µsec condensation clicks, presented at 80 dB nHL through
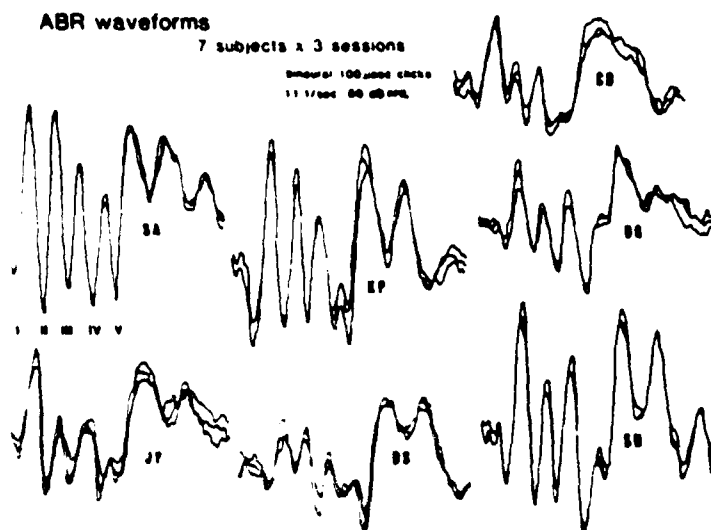
ABR waveforms



Fig. 1. A sampling of auditory brainstem responses (ABRs) for 7 young normal subjects [positive downward]. For each subject, waveforms from three sessions are shown superimposed; the five positive peaks under study are labelled in the traces for subject S. A. These between-subject and within-subject comparisons of the crude waveform profile corroborate earlier observations that: 1) there are clear individual differences to be seen in the ABR waveform, and 2) individual subjects' waveform shapes are quite consistent from session to session.

Telex 1470 earphones with MX-41/AR cushions. ABRs were processed using a Nicolet CA-1000 system sampling once every 20 μsec; records were stored on magnetic disk and analysed off-line. Subjects reclined with eyes closed. Stimuli were presented at a rate of 11.1 clicks per second; 2000 responses were averaged using a time window of 10 msec post-stimulus onset, and a filter setting of 150 to 3000 Hz (−3dB) with a 6 dB/octave roll-off. The artifact rejection criterion was 20 μvolts peak-to-peak.

## RESULTS

Samples of ABRs observed in the 7 subjects for the initial test session and two re-test sessions are shown in Fig. 1. These records are similar to those published previously, and illustrate that there are differences among ABRs from subject to subject, and that within-subject replicability of the waveforms is good.

In order to perform quantitative comparisons of such waveform characteristics, we obtained the latency in msec for each ABR peak for each presentation mode for each subject in each session. These latency values were then averaged across subjects across sessions to obtain an overall between-subject mean and standard deviation of the latency of each peak. The latencies were also averaged within

each subject across sessions to obtain a within-subject mean and standard deviation of the latency of each peak for each subject. To compare the relative variability of peaks having different absolute latencies, we calculated the coefficient of variation (Cv) for the different subject and session combinations. Since the Cv is simply the mean divided by the standard deviation, it varies inversely with standard deviation, and thus can be taken as a measure of stability: the smaller the standard deviation, i.e., the more stable the response, the larger will be the Cv. Table I presents comparisons of between-subject vs. within-subject calculations of means and Cvs for ABR peak latencies.

Cv values are displayed in Fig. 2. At the bottom of the figure are shown between-subject 'Cv profiles' for the five ABR peaks, for the group of 7 subjects across the eight test sessions. Responses for left-ear stimulation are indicated by open circles, for right-ear stimulation with filled circles, and for binaural stimulation by triangles. Analysis of variance indicates that for these between-subjects comparisons, there are significant differences in response latency as a function of peak ($F = 23.22$ [signif. $F$ at a confidence level of 0.01 with 4,28 degrees of freedom = 4.07]), but no differences due

Table 1. *Comparison of between-subject and within-subject calculations of actual values and coefficient of variation for peak latencies of five brainstem vertex-positive peaks*

Values given for between-subject calculations represent the mean (actual value or Cv) of 4 or 8 sessions' averages, where each session's average is calculated over 7 subjects. Values given for within-subject calculations represent the mean (actual value or Cv) of 7 subjects' averages, where each subject's average is calculated over 4 or 8 sessions

| | | P I | | | P II | | | P III | | | P IV | | | P V | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | R | L | B | R | L | B | R | L | B | R | L | B | R | L | B |
| *Between-subject* | | | | | | | | | | | | | | | | |
| Sessions | X̄ | 1.56 | 1.49 | 1.52 | 2.73 | 2.69 | 2.74 | 3.71 | 3.71 | 3.68 | 4.75 | 4.59 | 4.81 | 5.57 | 5.58 | 5.62 |
| 1–4 | Cv | 12 | 15 | 14 | 34 | 40 | 37 | 41 | 38 | 35 | 34 | 47 | 46 | 30 | 31 | 27 |
| Sessions | X̄ | 1.54 | 1.49 | 1.51 | 2.71 | 2.71 | 2.75 | 3.69 | 3.70 | 3.69 | 4.84 | 4.74 | 4.83 | 5.58 | 5.60 | 5.61 |
| 5–8 | Cv | 13 | 16 | 14 | 36 | 40 | 38 | 39 | 39 | 44 | 51 | 60 | 45 | 30 | 28 | 29 |
| Sessions | X̄ | 1.55 | 1.49 | 1.51 | 2.72 | 2.70 | 2.74 | 3.70 | 3.70 | 3.69 | 4.79 | 4.82 | 4.82 | 5.58 | 5.59 | 5.62 |
| 1–8 | Cv | 12 | 15 | 14 | 38 | 41 | 38 | 40 | 39 | 39 | 42 | 54 | 45 | 30 | 29 | 28 |
| *Within-subject* | | | | | | | | | | | | | | | | |
| Sessions | X̄ | 1.56 | 1.49 | 1.52 | 2.72 | 2.69 | 2.74 | 3.71 | 3.71 | 3.70 | 4.85 | 4.89 | 4.81 | 5.57 | 5.58 | 5.62 |
| 1–4 | Cv | 37 | 40 | 45 | 85 | 90 | 97 | 110 | 96 | 157 | 81 | 95 | 143 | 116 | 91 | 136 |
| Sessions | X̄ | 1.54 | 1.50 | 1.51 | 2.70 | 2.71 | 2.75 | 3.70 | 3.70 | 3.69 | 4.85 | 4.90 | 4.83 | 5.58 | 5.60 | 5.61 |
| 5–8 | Cv | 38 | 42 | 45 | 62 | 87 | 82 | 89 | 129 | 130 | 88 | 90 | 129 | 129 | 129 | 165 |
| Sessions | X̄ | 1.55 | 1.49 | 1.52 | 2.71 | 2.70 | 2.74 | 3.71 | 3.70 | 3.70 | 4.85 | 4.89 | 4.82 | 5.58 | 5.59 | 5.62 |
| 1–8 | Cv | 36 | 39 | 44 | 75 | 73 | 113 | 89 | 94 | 124 | 90 | 105 | 142 | 111 | 84 | 120 |



ABR latency Coefficient of variation
5 peaks brainstem
7 subjects x 8 sessions
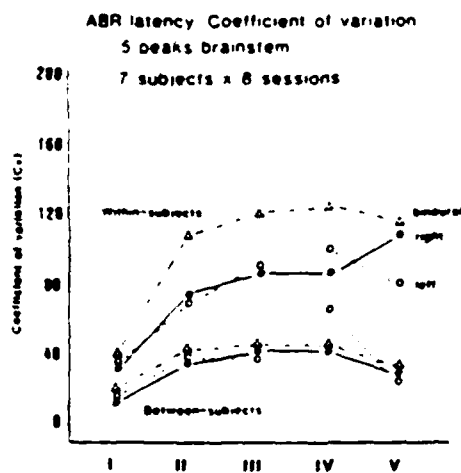
*Fig. 2.* Comparisons of between-subject 'Cv profiles' (bottom) with within-subject Cv profiles (top). (O), responses for left-ear stimulation; (●), for right-ear stimulation, and (*triangles*) for binaural. Each data point is based on fifty-six values (7 subjects times eight test sessions). Between-subject values are derived by: (*a*) calculating a between-subject Cv for each session across the 7 subjects, and then (*b*) taking the mean of these eight values (one for each session) to obtain a 'mean between-subject Cv' for each peak. Within-subject values for each peak are derived by: (*a*) calculating a within-subject Cv for each subject across the eight sessions, and then (*b*) taking the mean of these seven values (one for each subject) to obtain a 'mean within-subject Cv' for each peak.

to ear stimulated ($F = 1.16$ [signif. $F$ at a confidence level of 0.05 with 2,14 degrees of freedom = 3.74]), and no significant interaction between peak and ear ($F = 1.00$ [signif. $F$ at a confidence level of 0.05 with 8,56 degrees of freedom = 2.21]).

Cv can also be calculated for within-subject variability. For each peak for each ear condition (right, left, binaural), we obtained each subject's Cv over the eight test sessions, and then took the average of these values over the 7 subjects to obtain a 'mean within-subject Cv', for each peak of the ABR; values are included in Table I. Within-subject Cv profiles are shown at the top of Fig. 2. Note that in all cases, the stability of within-subject comparisons is greater than between-subject comparisons, i.e., subjects are more like themselves than like each other, as suggested by the recordings of Fig. 1.

Analysis of variance for the within-subjects Cv data indicates that there are significant differences in response latency as a function of both peak ($F = 9.12$ [signif. $F$ at a confidence level of 0.01 with 4,24 degrees of freedom = 4.22]) and ear of presentation ($F = 5.13$ [signif. $F$ at a confidence level of 0.05 with 2,12 degrees of freedom = 3.89]). The interaction between peak and ear was not significant ($F = 0.06$ [signif. $F$ at a confidence level of 0.05 with 8,48 degrees of freedom = 2.14]).

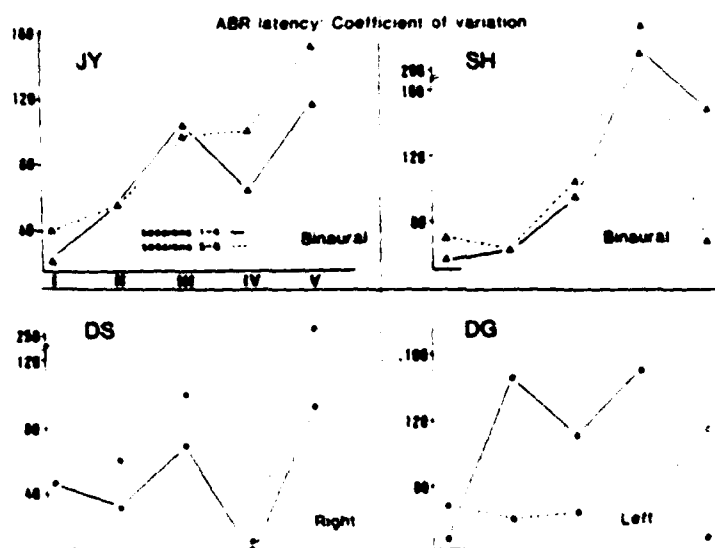Fig. 3 shows selected Cv curves for 4 individuals

Fig. 3. Selected within-subject ABR Cv profiles for 4 individuals. Values connected by solid lines are derived from the first four test sessions; values connected by broken lines are derived from the second four test sessions. Two sets of binaural responses are shown (subjects J. Y. and S. H.), one set of right-ear responses (D. S.), and one set of left-ear responses (D. G.).

comparing the first four sessions with the second four. There are clear individual differences, according to the subject tested and the ear stimulated, in these patterns. Of the 21 such patterns (7 subjects ×3 presentation modes), 7 show replicability as good as those shown here for J. Y., D. S., and S. H.; another 6 show departures from good replicability due to *increased* stability in one or in two peaks. Only eight patterns, including D. G.'s left-ear profile, fail to replicate because of *decreased* stability at one or more peaks. This pattern replicability and typical increases in the stability of individual responses over time lead us to believe tha these patterns are reflections of true individual differences in brainstem response, and are not random.

Finally, we considered whether overall Cv profiles by ear of presentation would show similar replicability over time. Fig. 4 shows within-subject Cv profiles for right, left, and binaural stimuli, comparing results of the first four test sessions versus the second four. Analysis of variance showed significant differences in peak latencies as a function of peak ($F = 24.61$ [signif. $F$ at a confidence level of 0.01 with 4,8 degrees of freedom $= 7.01$]), but none for time ($F = 1.48$ [signif. $F$ at a confidence level of 0.05 with 1,4 degrees of freedom $= 7.71$). Note that in both halves of the test series for this group of subjects, right and binaural responses increase in stability through peak III, fall at IV, and rise again

at V; and stability at V appears to *increase over time* for *both* presentations. In contrast, left-ear responses peak at *IV* and *fall* at V, with this pattern becoming even more pronounced in the second half of testing.

## DISCUSSION

Interpretations of auditory evoked responses collected in a clinical setting are clearly dependent on knowledge of the range of normal variability. A number of reports have recently documented the details of this variability, for between-subject comparisons. However, although data regarding within-subject variability are also of clinical interest, e.g., for following the time-course of progressive disorders such as multiple sclerosis, or of recovery from surgery, little is known regarding the changes in AER responses in one subject over time. Our data provide a first step toward describing such variability. The results corroborate the evidence apparent in ABR tracings, that individuals are more like themselves than like each other—not a surprising result, but demonstrated here quantitatively. Second, these data suggest that, at least for ABRs, there are clear reflections of individual differences and of mode of presentation to be discerned in the Cv profiles describing the relative stability of different peaks in the evoked response waveform.

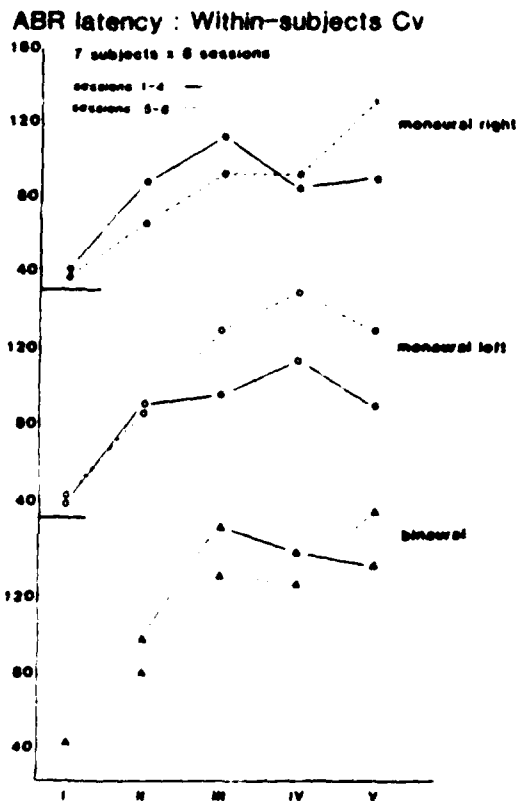At present it is not clear what accounts for the

## ABR latency : Within-subjects Cv



*Fig. 4* Mean within-subject ABR Cv profiles [calculation as described in Fig. 2] compared for ear of presentation: ○, left-ear responses; ●, right-ear responses; and binaural responses (*triangles*). Values connected by solid lines are derived from the first four test sessions; values connected with broken lines are derived from the second four test sessions.

individual differences, or for the differences correlated with ear of presentation. Further conclusions regarding the significance of such differences await analysis of other response parameters, such as interpeak latencies, relative amplitudes, and comparisons of monaural and binaural patterns of response. as well as chronological studies of the patterns of evoked responses followed over a longer period of time.

These results have implications for both hearing research and application. First, they suggest that non-invasive measures of evoked responses may have more potential for revealing details of central auditory function than has been previously realized. Based on the patterns of response revealed

here, it seems possible that profiles of individual response variability could be used to study the physiological organization of individual auditory nervous systems. For example, it is interesting that significant effects due to ear of presentation can be seen in the within-subject but not in the between-subject comparisons (cf Fig. 2).

Second, knowledge regarding the details of individual variability could be important for certain clinical applications. Patterns of response could be used in studies of developmental changes within the central auditory nervous system, and for charting deviations from a normal developmental path. Applications such as monitoring the course of progressive disorders, or in tracking recovery from brainstem surgery, should find baseline measurements regarding patterns of individual variability useful.

## REFERENCES

Ainslie. P. J. & Boston. J. R. 1980. Comparison of brain stem auditory evoked potentials for monaural and binaural stimuli. *Electroenceph Clin Neurophys 49*. 291.

Allison. T.. Wood. C. C. & Goff. W. R. 1983. Brain stem auditory, pattern-reversal visual, and short-latency somatosensory evoked potentials: latencies in relation to age. sex. and brain and body size. *Electroenceph Clin Neurophys 55*. 619.

Aunon. J. I. & Cantor. F. K. 1977 VEP and AEP variability: interlaboratory vs. intralaboratory and intersession vs. intrasession variability. *Electroenceph Clin Neurophys 42*. 705.

Beagley. H. A. & Sheldrake. J. B. 1978. Differences in brainstem response latency with age and sex. *Brit J Audiol 12*. 69.

Bergholtz. L. 1981. Normative data in clinical ABR. *Scand Audiol.* Suppl. *13*. 75.

Blegvad. B. 1975. Binaural summation of surface-recorded electrocochleographic responses; normal-hearing subjects. *Scand Audiol 4*. 233.

Chiappa. K. H.. Gladstone K. J. & Young. R. R. 1979. Brain stem auditory evoked responses; studies of waveform variations in 50 normal human subjects. *Arch Neurol 36*. 81.

Dobie. R. A & Norton. S. J. 1980. Binaural interaction in human auditory evoked potentials. *Electroenceph Clin Neurophys 49*. 303.

Edwards, R., Buchwald, J. S., Tanguay, P. E. & Schwafel, J. A. 1982. Sources of variability in auditory brain stem evoked potential measures over time. *Electroenceph Clin Neurophys 53*, 125.

Goodin, D. S., Squires, K. C., Henderson, B. H. & Starr, A. 1978. Age-related variations in evoked potentials to auditory stimuli in normal human subjects. *Electroenceph Clin Neurophys 44*, 447.

Kendall, J. P & Lawes, I. N. C. 1978. The clinical reliability of brain stem auditory evoked responses. *Brit J Audiol 12*, 23.

Kjaer, M. 1979. Differences of latencies and amplitudes of brain stem evoked potentials in subgroups of a normal material. *Acta Neurol Sçand 59*, 72.

Kjaer, M. 1980. Recognizability of brain stem auditory evoked potential components. *Acta Neurol Scand 62*, 20.

Levine, R. A. & McGaffigan, P. M. 1983 Right–left asymmetries in the human brain stem: auditory evoked potentials. *Electroenceph Clin Neurophys 55*, 532.

Owen, J. H. & Matsusaka, E. K. 1982. Influence of test parameter combinations on the auditory evoked late response. *Electroenceph Clin Neurophys 53*, 329.

Rosenhamer, H. J., Lindström, B. & Lundborg, T. 1978. On the use of click-evoked electric brainstem responses in audiological diagnosis. I. The variability of the normal response. *Scand Audiol 7*, 193.

Rosenhamer, H. & Holmkvist, C. 1982. Bilaterally recorded auditory brainstem responses to monaural stimulation: latency and amplitude differences in normal subjects. *Scand Audiol 11*, 197.

Rowe, M. J., III 1978. Normal variability of the brain stem auditory evoked response in young and old adult subjects. *Electroenceph Clin Neurophys 44*, 459.

Spreng, M. 1980. Variability of the early (0–10 ms) auditory evoked extracranial components. *Scand Audiol, Suppl. 11*, 79.

Stockard, J. E., Stockard, J. J., Westmoreland, B. F. & Corfits, J. L. 1979. Brainstem auditory-evoked responses; normal variation as a function of stimulus and subject characteristics. *Arch Neurol 36*, 823.

Stockard, J. E., Stockard, J. J. & Coen, R. W. 1983. Auditory brain stem response variability in infants. *Ear and Hearing 4*, 11.

Thornton, A. R. D. 1975. Bilaterally recorded early acoustic responses. *Scand Audiol 4*, 173.

## ANNOUNCEMENTS

*The American Medical Writers Association presents 1986 Annual Conference. Getting to the Heart of Things in San Francisco.* October 22–25, 1986. Sherton Palace Hotel, San Francisco, CA.

Journal Editors; Writers; Media and Freelance Personnel; Audiovisual Writers and Producers; Advertising, Public Relations and Pharmaceutical Professionals; Physicians, Dentists, Nurses and Other Health Care Professionals; Translators; Indexers; Hospital Administrators; Medical Communicators.

Update your knowledge through top notch workshops, stimulating general sessions, networking breakfasts, award lunches, section meetings, and varied social activities.

For further information, contact: Lil Sablack, Executive Director, American Medical Writers Association, 5262 River Rd., Suite 410, Bethesda, MD 20816, USA tel. (301) 986 9119.

*The Departments of Otolaryngology and Pediatrics, Children's Hospital of Pittsburgh present The 13th Annual Symposium: Ear, Nose and Throat Diseases in Children: A 1986 Update.* December 3–7, 1986, The Breakers, Palm Beach, Florida.

Otolaryngologists, Pediatricians, and other Health Care Professionals: Update your knowledge through scientific

presentations by University of Pittsburgh School of Medicine faculty and distinguished invited speakers, small-group workshops, and informal discussions.

CME Credits: 17 hours Tuition: $275 Physicians, $200 Residents.

For further information, contact: Dept. of Otolaryngology, Children's Hospital of Pittsburgh, 125 De Soto St., Pgh., Pa. 15213 USA tel (412) 647-6215.

*American Auditory Society.* The 14th annual meeting of the American Auditory Society will be held in Chicago, Illinois on Monday, September 21, 1987. The meeting is scheduled in conjunction with the annual meeting of the American Academy of Otolaryngology—Head and Neck Surgery.

*4th International Meeting on Low Frequency Noise & Vibration,* to be held in Umeå on June 9th, 10th & 11th 1987 at University of Umeå, Humanisthuset, Sweden.

*Call for Papers.* The Organizing Committee welcomes contributions on any of the topics listed opposite. An abstract (with title) of about 200 words should be submitted to Dr. W. Tempest, Multi-Science Publishing Co. Ltd., 107 High Street, Brentwood, Essex CM14 4RX, England. To arrive before 30th November 1986.

Although studies with synthetic speech have suggested that voice-onset time (VOT) provides an important cue to the identification of stop consonants, the variability of VOT in real speech has not been studied to any extent, either for several productions by one talker, or for several talkers. More information regarding the variability of this and other acoustic cues in speech may shed light on how listeners learn to generalize the multitude of acoustical patterns of speech into a few phonemic categories.

SLIDE 1

For this study, sets of 6 stop consonants were recorded in nonsense di-syllables, with each stop preceding the vowel /a/. Using these lists we have recorded productions from 5 female and 1 male English talker, 3 female Japanese talkers, and 3 female Spanish talkers.

SLIDE 2

A list of real words was used to elicit stop consonants from two female Navajo speakers, since they could not interpret written nonsense syllables as Navajo phonemes. Five Navajo stops were recorded: b, d, g, t-plus-glottal-stop, and k-plus-glottal stop, each preceding the vowel /a/.

Recordings were done in an anechoic chamber. Each talker read the list of syllables (or words) six times, and productions were digitized and stored on video tape. A Kay 7800 digital sonograph was then used to produce waveforms for each sound.

SLIDE 3

Sets of waveforms were then prepared comparing the VOT portion of each of the six instances of each consonant produced by each talker. These sets of six were examined to determine where in each talker's waveform voice onset occurred. Our criteria for judging the point of voice onset were based solely on the shape of the waveform, as we sought a point where the noise of the consonant began to change into a more articulated periodic form presumed to reflect the onset of vocal-fold action.

**SLIDE 4**

Occasionally, such sets of these productions revealed sequential VOT series. For example, this set for this talker shows regular increments as VOT duration increases from bilabial through velar voiced stop consonants, and again from bilabial through velar voiceless stops. However, our recordings indicate that this degree of regularity is the exception rather than the rule.

**SLIDE 5**

Here we see the complete series of six repetitions of the six stops by this same talker. The bar graph of set #1 represents the waveforms we just saw. Note however that in most sets, one distinction or another is poorly represented by VOT: for example, in #3, /b/ and /d/ are almost identical in VOT, as are /t/ and /k/ in #4. In some cases, the expected differences in VOTs are reversed: for example, in #5, /d/ is shorter than /b/ and /k/ is shorter than /t/. The one constant for this talker seems to be the distinction between voiced and voiceless; for all six repetitions, the voiced VOTs are shorter than 30 ms, while the voiceless VOTs are consistently longer than 50 ms.

**SLIDE 6**

The next few slides present summaries of such measurements, in terms of mean and range of VOT for the 6 repetitions of each stop by each talker. As we will see, for all 6 English speakers and for the 2 Navajo talkers, VOTs provide somewhat ambiguous cues to place of articulation within voiced or voiceless sets, but for all these speakers, VOTs show a clear and consistent distinction between voiced versus voiceless consonants. These are the first three English female talkers;

**SLIDE 7**

Here are the other two English females and one male talker (DH);

**SLIDE 8**

Here are the two Navajos.

**SLIDE 9**

For the 3 Spanish and 3 Japanese talkers, mean positive VOTs do not distinguish between voiced and voiceless stops. These are the Spanish results, and the next slide--

**SLIDE 10**

shows the results for the Japanese talkers. For some of these subjects, there appear to be longer VOTs for the velar stops.

The relative variability of VOTs, suggested in these slides by the range measurements, can be expressed directly--

SLIDE 11

for example, in terms of the standard deviation considered as a percentage of the mean. This slide shows results of such a calculation for each of the six stops produced by the first 3 English talkers. Voiced stops are indicated with closed circles, voiceless stops with open circles.

For example, for JL, the distribution of VOTs for her six productions of /b/ has a standard deviation that is approximately 30% of the mean, while her /p/ VOT distribution is much tighter, with a standard deviation that is only 10% of the mean.

Calculations of this measure for all the talkers show clear individual differences, but as we will see, there also seem to be agreements in the patterns of variability shown by different talkers. For example, the pairs of voiced and voiceless curves for English talkers JL and CB resemble each other,

SLIDE 12

as do the curve pairs for NF and DH.

SLIDE 13

Navajo talker JJ's voiced-stop curve looks like the voiced-stop curve of English talker CB, while Navajo talker BM's voiced-stop curve looks like the analogous curve for English talker AS. We will see direct comparisons of these patterns in a moment.

SLIDE 14

Here, VOT variability patterns for the 3 Spanish subjects not only resemble each other, but are similar to those shown by 2 English talkers.

SLIDE 15

And for the Japanese, two of these sets of curves resemble English patterns. The results for Japanese talker KH represent a unique VOT variability pattern in these data.

SLIDE 16

These agreements in the form of VOT variability patterns that seem to cut across language boundaries are summarized in the last slide. Here our variability measure for VOT is shown as a function of the consonant being produced, with language of the talker as a parameter. Pattern Type 1 at the top is shown by a total of 6 of these 14 talkers: all 3 Spanish, 2 English, and 1 Navajo.

Pattern Type 2 is shown by two English, 1 Navajo, and 1 Japanese, and Pattern Type 3 is shown by two English and 1 Japanese. Type 4 is the unique pattern shown only by Japanese talker KH.

Admittedly this description of variability patterns in VOTs is extremely preliminary. It is possible that with further testing, the patterns we have observed for the different talkers would change. However, the observation that--even within this small population of talkers--individuals fall into groups, and, that the same patterns occur in several languages, suggests that the different variability patterns defined here may in fact represent a limited number of styles of VOT production, that may be language-independent.
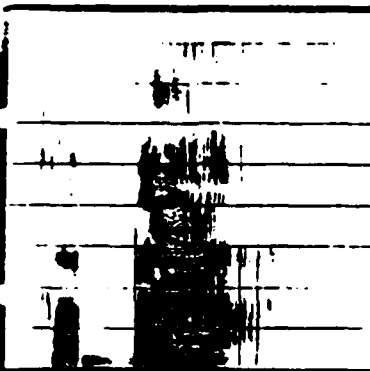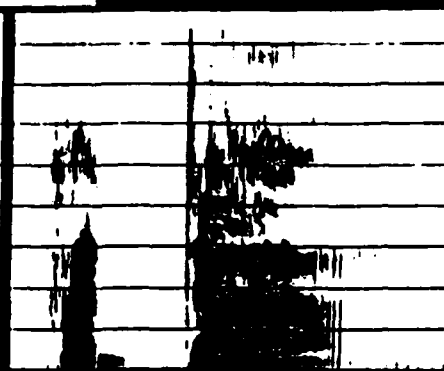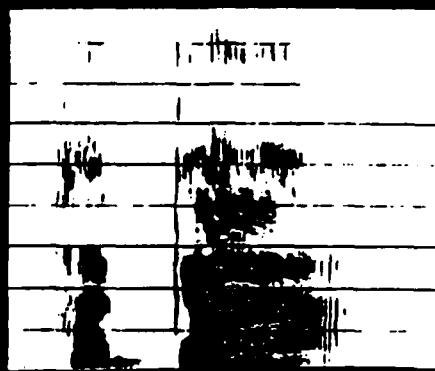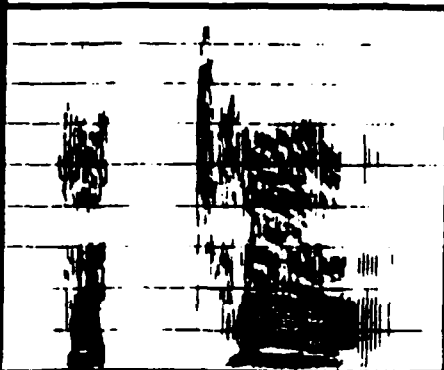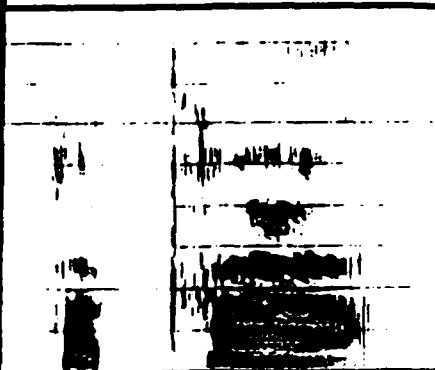
An obvious next step is to study these same utterances further to determine whether similar variability measurements of cues other than VOT will group the listeners in the same or in complementary ways. For example, we might ask whether the variability of VOTs on the one hand, and of burst spectra on the other, are inversely related: perhaps talkers from the Type 1 group, with very stable voiceless VOTs, allow their voiceless burst spectra to vary more than do talkers in the Type 2 group, with relatively unstable voiceless VOTs.

Thus our preliminary conclusion from this work is: that acoustic-cue tradeoffs used to signal different phonemes may be described in terms of the variability as well as the absolute values of the cues--and that there may be several ways to organize the structure of these tradeoffs. Further work with this data base, and with productions from new talkers, are planned to help us test these hypotheses.

Thus it may be that in speech perception, learning to analyze the details of the speech waveform is only the first--and easiest--step: the listener must then go on to learn not only the constellations of cues used by each talker for each phoneme, but also to analyze and generalize across the different patterns of cue interactions used by different talkers to signal the same phonemes.
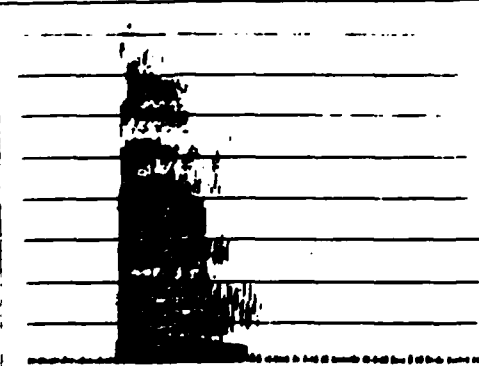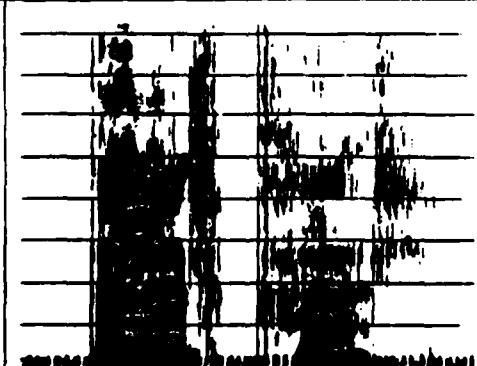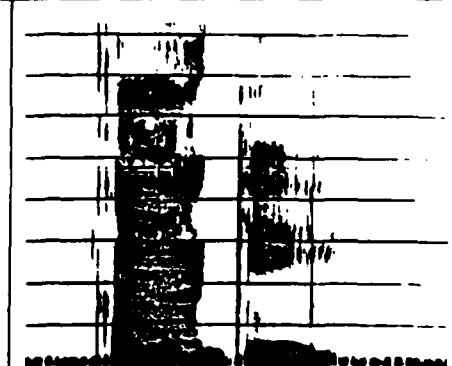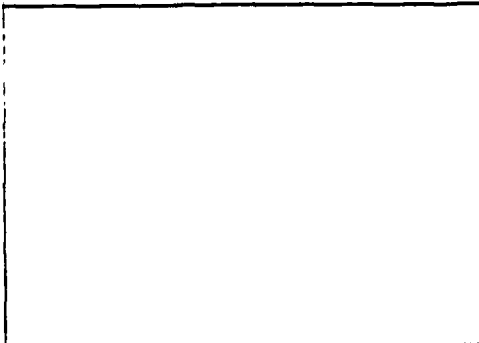

*******
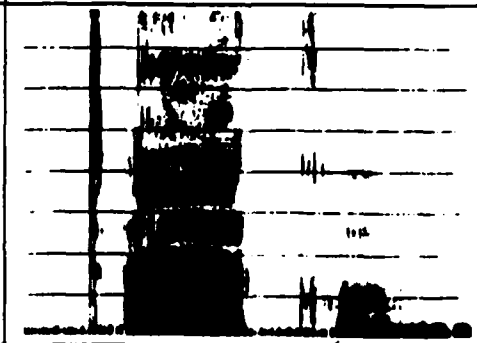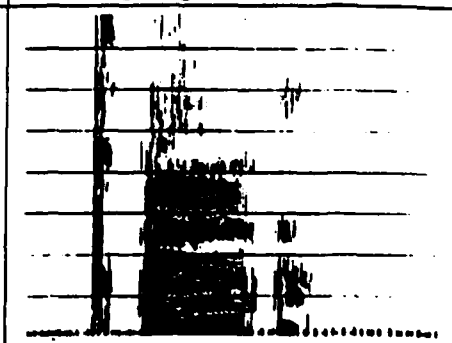
**English: 6 stop CV syllables**

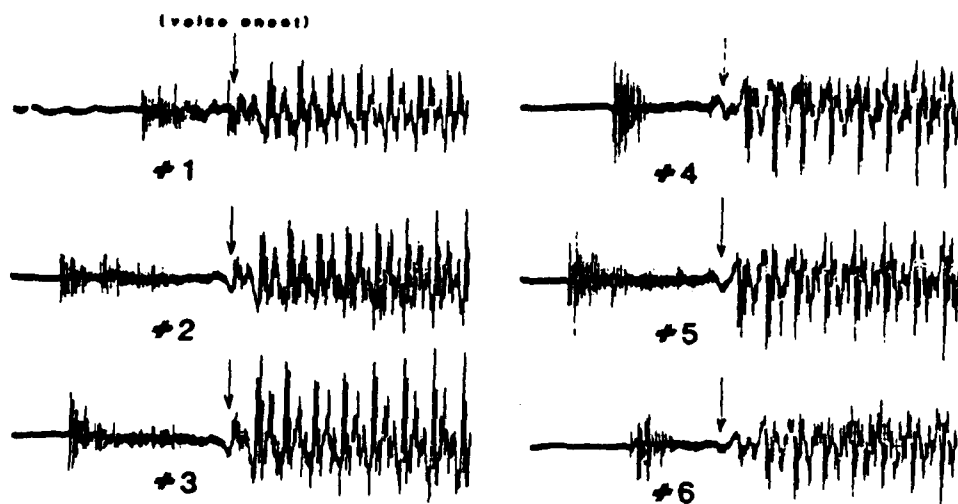

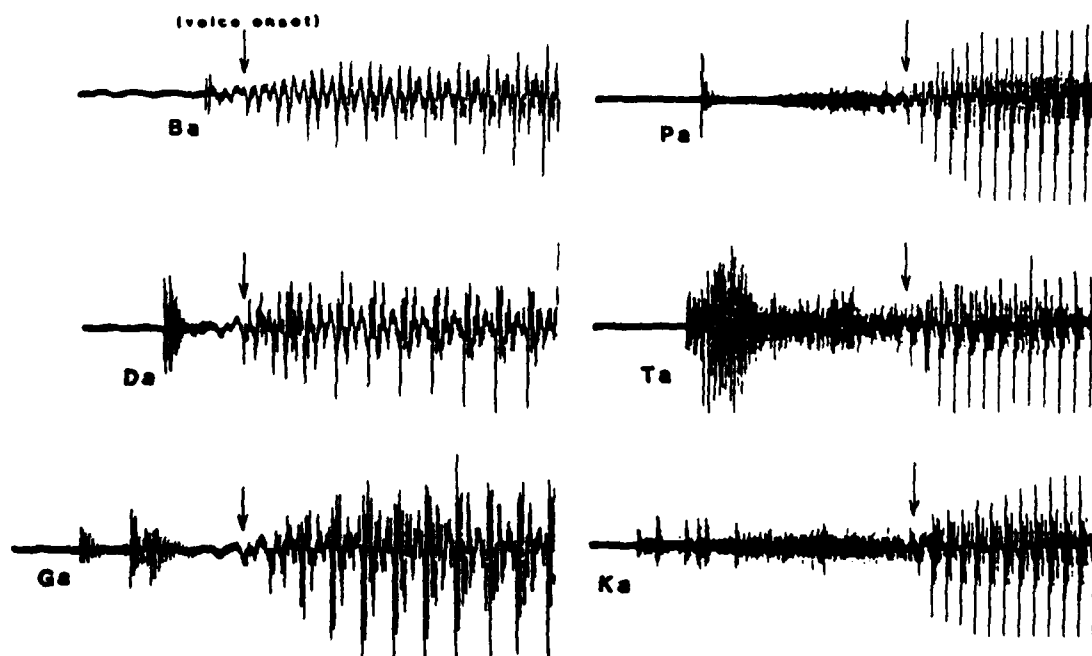| ha- Ba | Da | Ga |
| Pa | Ta | Ka |

**Navajo: 5 stop-initial words**



| bááh (bread) | dáádílkał (door) | gáagi (crow) |

# /Ga/: 6 tokens (English female: NP)
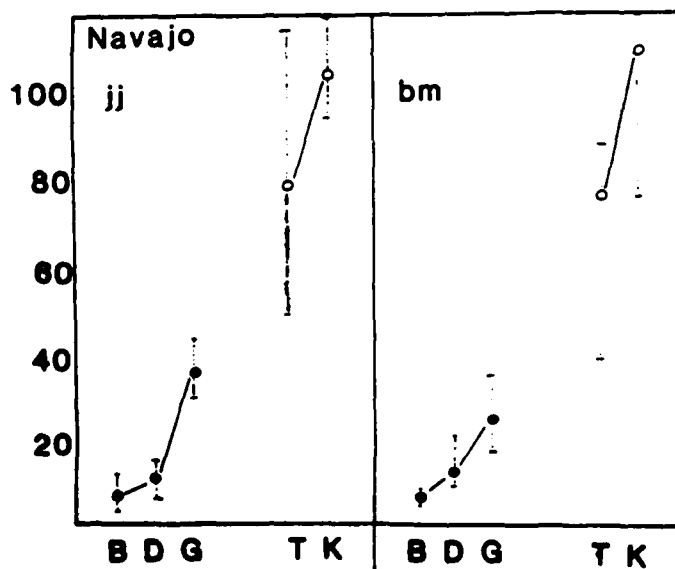


## Stop CV waveforms: burst to voice onset

talker: English f. (JL)
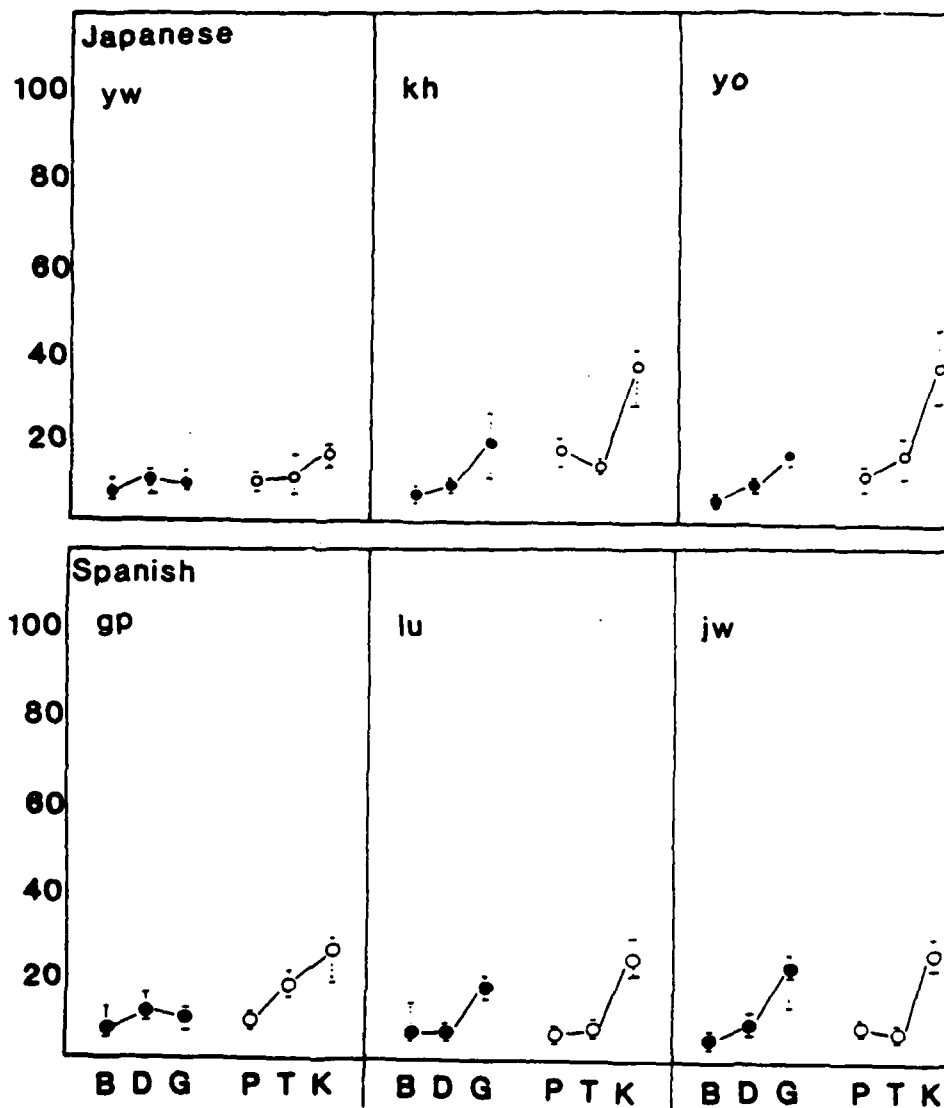
Six productions of sets of stop CVs (English f.:JL)

# VOT : mean and range in msec
## (6 tokens)



English: jl, cb, np, as, nf, dh

Navajo: jj, bm

B D G   P T K
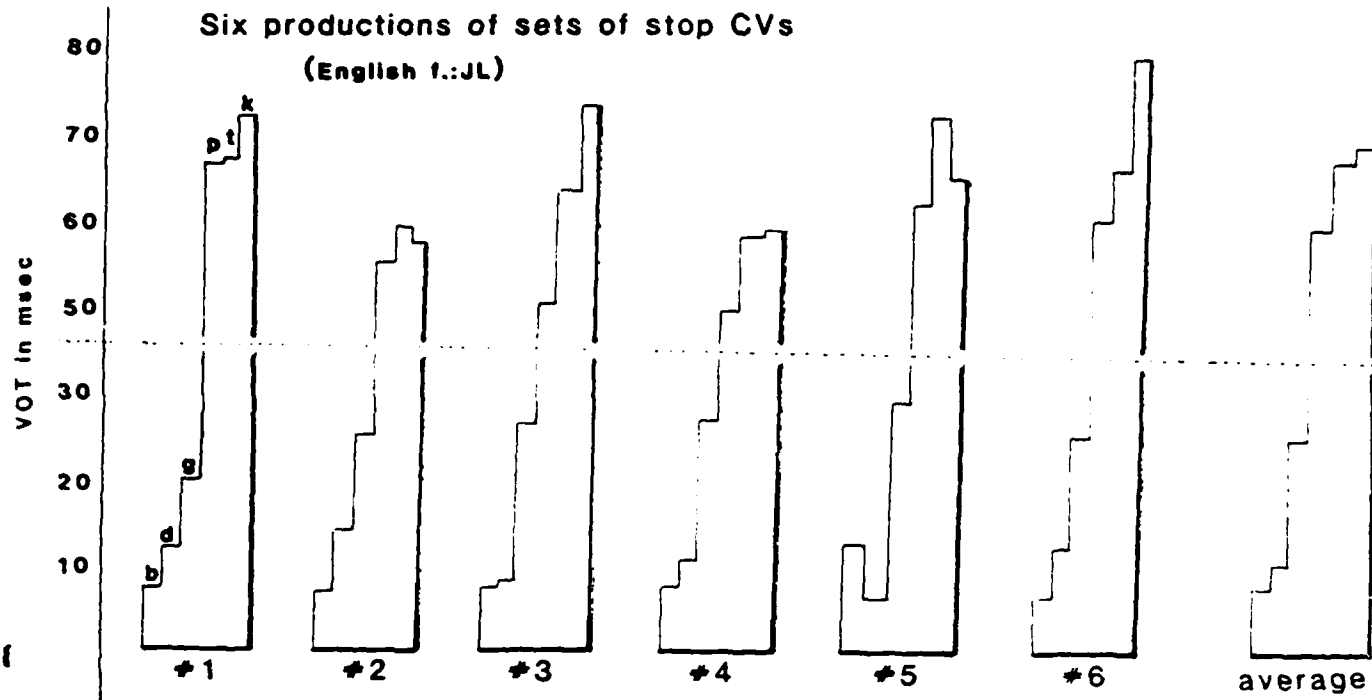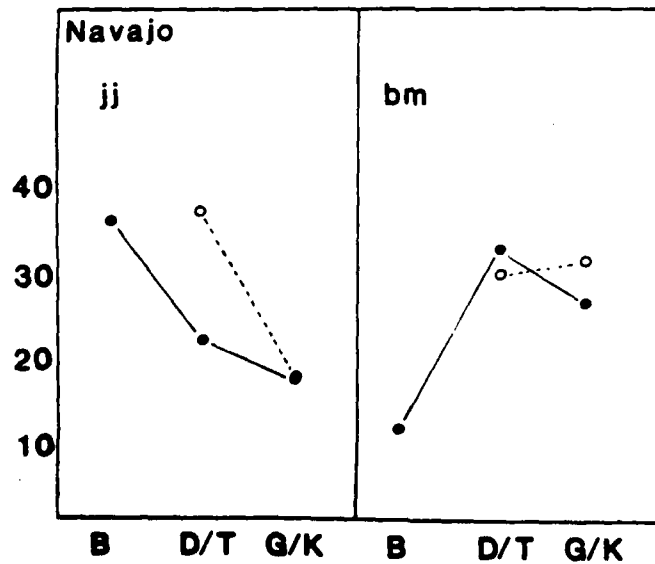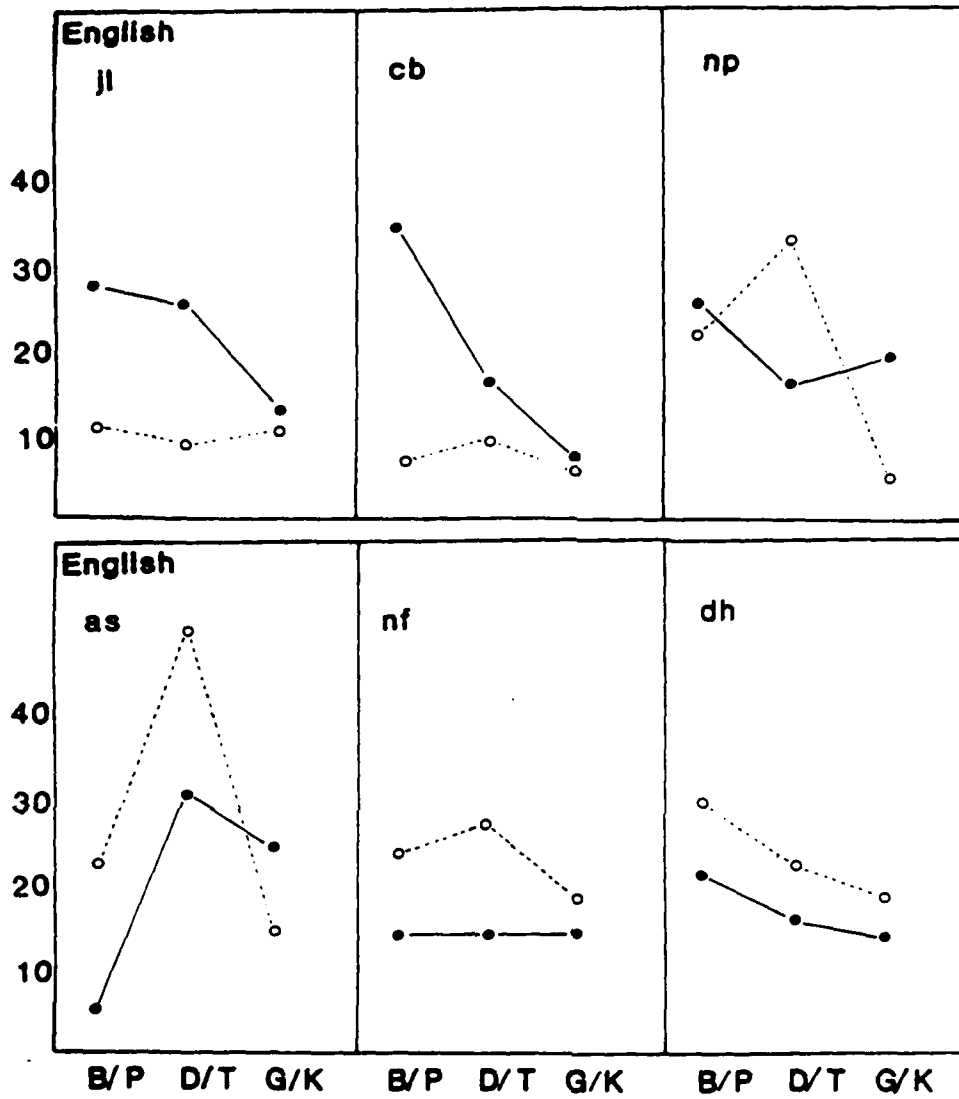
# VOT : mean and range in msec
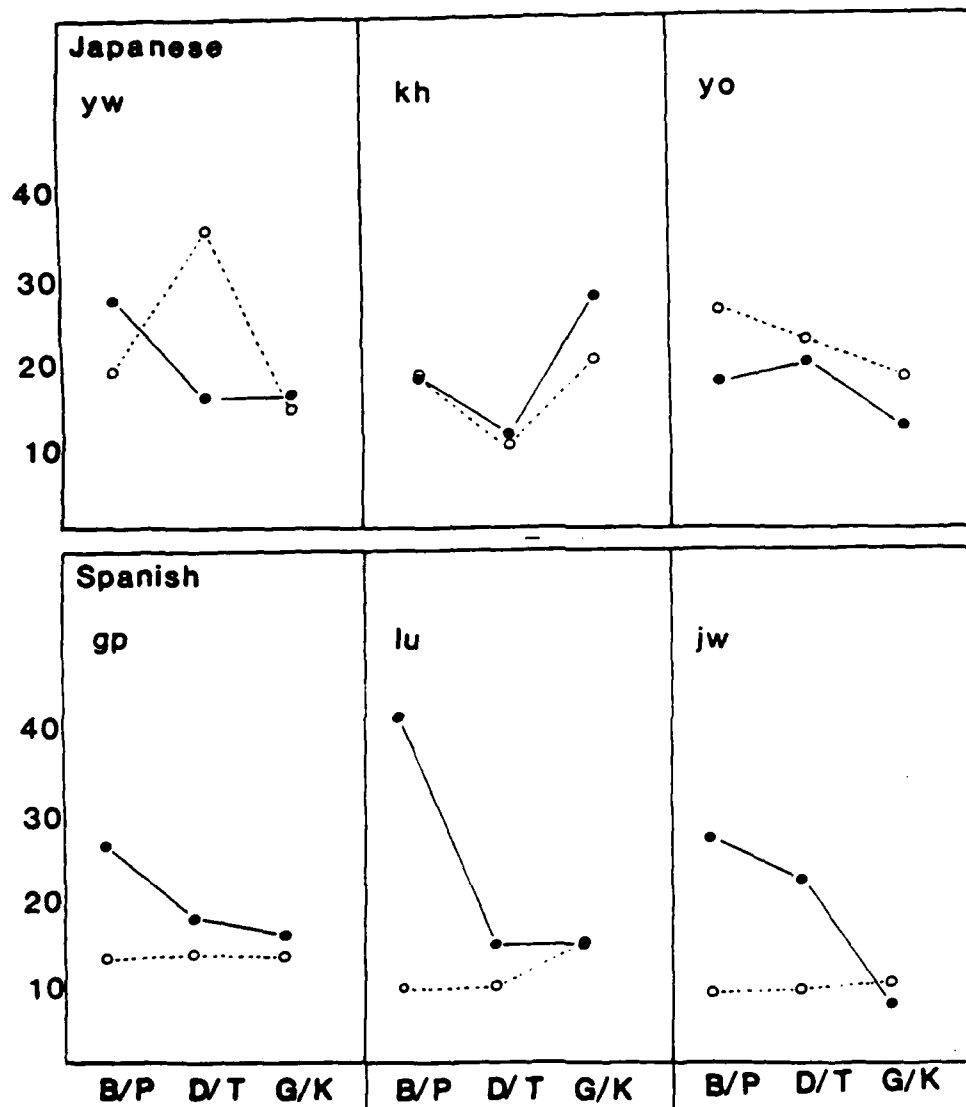## (6 tokens)
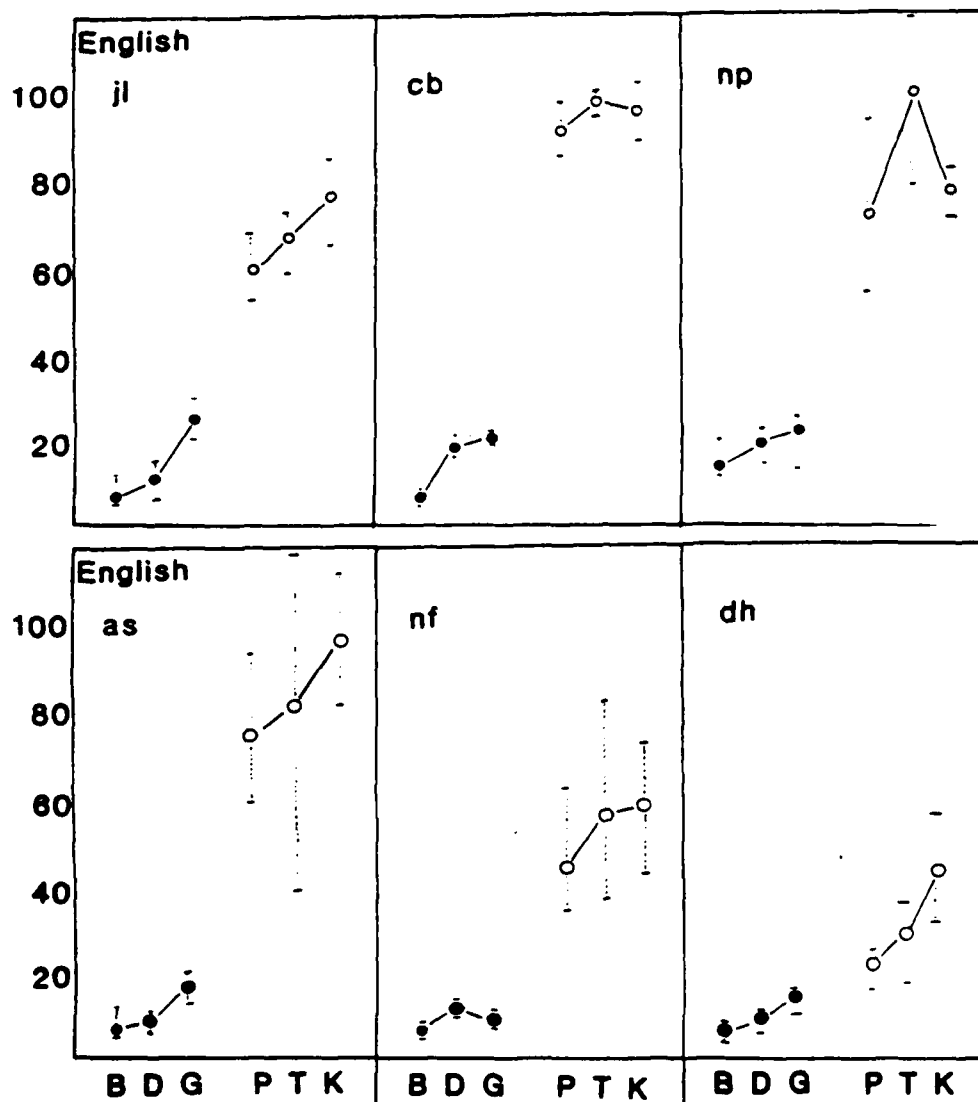
# VOT : s.d. as % of mean
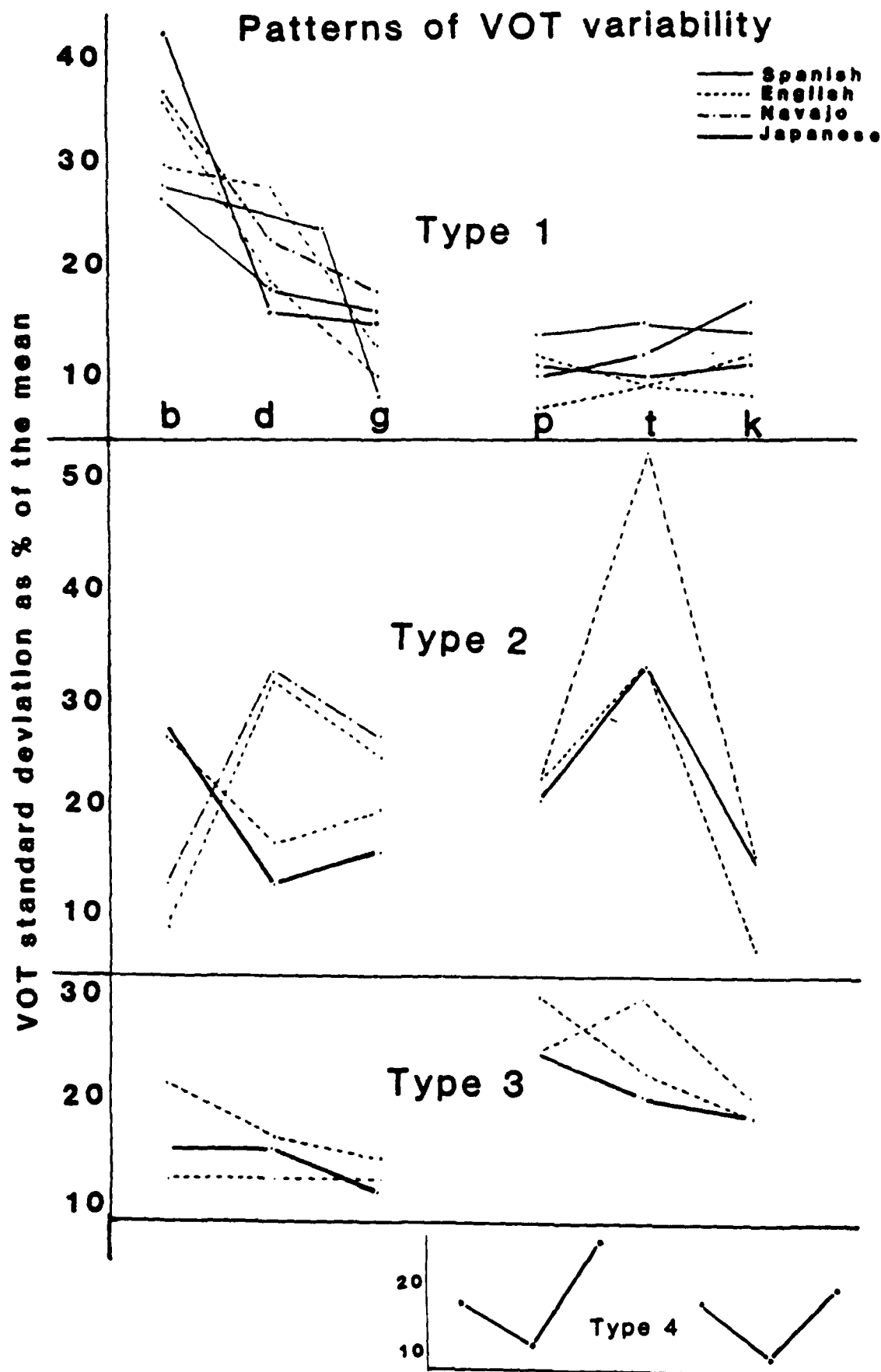
●—● voiced      o----o voiceless

# VOT : s.d. as % of mean

●—●voiced  ○---○voiceless

Patterns of VOT variability

Lauter JL and RG Karzon (1987) Individual differences in auditory-evoked potentials: Variability of middle-latency responses, including comparisons with brainstem AEPs. Presented to Acoustical Society of America, Indianapolis, May 1987. Abstract: JASA 81: S8.

As several investigators have observed, the waveform of the auditory evoked potential, even at the brainstem level, is not the same when compared from subject to subject (Slide 1). As can be seen in these tracings, taken from our first study reported here in 1985 on ABRs collected from 7 subjects in a series of 8 weekly sessions, the 7 different subjects show 7 different waveforms. And yet, as the superimposed tracings from three weekly sessions for each subject show, each individual's waveform replicates very well. In a series of reports to this society, we have described our attempts to quantify these observations in both ABRs and cortical responses, seeking to compare the degree of difference between individuals with the degree of similarity when individuals are compared with themselves.

We have found that a simple measure of relative variability, comparing waveform parameters such as peak latency and amplitude, can distinguish clearly between inter- and intra-subject consistency in EP waveforms. We have also found that variability of waveform parameters is sensitive to several characteristics of waveforms that cannot be distinguished in plots of the parameters themselves. (2)

For example, in the next slide, we compare measures based on ABR peak latency in msec (on the left) with our variability measure for peak latency on the right. In the panel on the left, peak latency in msec is plotted for each of five vertex-positive brainstem peaks, with one line each for right-ear, left-ear, and binaural presentation. Values are averaged over 7 subjects over 8 sessions each. Note that absolute latency shows no differences due to ear-of-presentation.

In contrast, the panel on the right is based on variability of peak latency . We call this measure the Coefficient of Stability, or Cs. It is calculated by dividing the mean of a value by its standard deviation. In this graph, we can see that simply by combining mean and standard deviation in measures of ABR waveforms, peak latency can be shown to reflect differences according to between- and within-subject comparisons (that is, inter- vs. intra-subject consistency), as well as differences according to ear-of-presentation, particularly in the within-subject comparisons, shown in the top three curves.

In today's presentation, we report an extension of this method to the study of middle-latency responses (MLRs). (Next slide) We will be describing the MLR waveform in terms of three vertex-negative and two vertex-positive peaks. For these tests, 8 new subjects, 4 females and 4 males, were recruited and screened to have hearing thresholds in both ears better than 20 dB nHL. Each subject was scheduled for a series of 8 weekly test sessions, with each session held on the same day at the same time of day each week. Each session consisted of testing for MLRs, with right-ear, left-ear, and binaural presentation conditions, followed by right, left, and binaural ABR testing. Electrodes were placed at vertex, mastoid, and forehead,

with monaural responses referenced to ipsilateral mastoid, and binaural responses referenced to linked mastoids. For both test series, stimuli were 10-per-second, 100-usec condensation clicks presented at 80 dB nHL. ABRs were collected using a window of 10 msec, filters set at 150-3000 Hz, and 2000 sweeps. MLRs were collected using a 50-msec window, filters at 50 to 1550 Hz, and 1000 sweeps.

Overall results for both MLR absolute latency and latency stability measures are shown on the next slide. (4) On the left, latency is plotted in msec for the five MLR peaks. As in the ABR results seen previously, absolute latency values reflect no differences due to ear-of-presentation. In contrast, the panel on the right displays the results of calculating the latency Coefficient of Stability for each of the five MLR peaks. Note that although the separation between curves for subject groupings and for ear-of-presentation is not as dramatic as for the ABR data, there are still distinctions to be seen, particularly in the calculations for MLR peak No.

This graph suggests, in fact, that our stability measure reflects the transitional nature of peak No, which a number of researchers have ascribed to brainstem generators. To explore this further, we compared these MLR stability curves with the ABR curves collected for the (5) same 8 subjects (next slide). Note that although both between- and within-subject curves for all MLR peaks subsequent to No show very poor stability, i.e., all are within the range of the between-subjects curves for the ABR data, the within-subject values for No seem to be in-between those for ABR peak V and those for MLR peak Po.

The transitional nature of MLR peak No can be seen even more clearly if we turn to individual data. The next slide shows stability profiles for two subjects. (6) Comparisons are shown between stability values for ABR peak V and for the same subject's five MLR peaks, for right, left, and binaural presentations. For both, values for No are intermediate betwen those for ABRV and later MLR peaks. (7)

Individual stability profiles such as these also show good replicability. The next slide shows panels for 3 subjects, NN Right ear, CH Binaural, and WB Left, comparing profiles calculated for the first 4 sessions versus those for the second 4. These comparisons are typical of those we see in all our repeated-measures EP data, indicating that stability profiles for individual subjects most often either replicate exactly over two months time, or depart from replication because of increases in stability of the response at one or more peaks.

Finally, we might ask how these stability data for the MLR fit into the larger context of similar measures for all three auditory EP levels--ABR, MLR, and cortex. (Next slide). (8) In the top panel of this figure are shown data for our original 7 subjects comparing their stability profiles for ABR and cortical responses. Note that both the between- and within-subject curves for cortex are very low: both are within the range of the between-subject ABR curves. Then, in the lower panel, we have repeated these curves, and inserted the MLR stability profiles for the new set of subjects. Although this is only a preliminary comparison, since we currently lack measures at all three levels in the same subjects, a segregation in MLR peaks

revealed by stability profiles is clearly suggested: the within-subject values of peak No are more similar to those for the ABR peaks, while values for the later MLR peaks are more like those seen for cortex.

Current plans for further testing which will allow us to explore these and other phenomena related to repeated-measures evoked-potential testing include: *(9)*

1) longer time series, to determine whether the characteristics of individual stability profiles we have described tend to remain constant over time, vary randomly, or develop in a systematic way; and

2) try more condensed test schedules, to see whether the same amount of data collected over a shorter time period shows the same patterns;

3) combine repeated-measures tests for ABR, MLR, and cortical responses, in the same subjects, to compare stability at different levels in the same system;
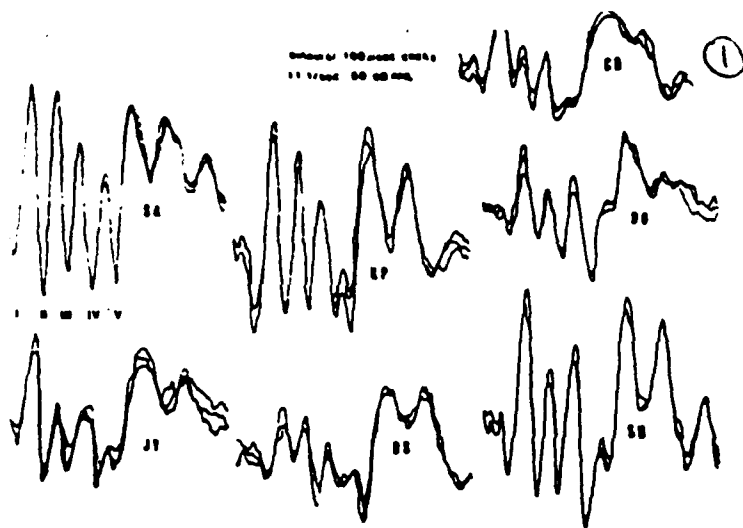
4) study subjects of different ages, to determine whether EP stability profiles can provide information regarding details of auditory development, including stages of degree of individual differentiation, asymmetrical response, and time course of development at different levels of the auditory CNS;

5) compare the individual asymmetries shown by our stability profiles with those measured by the BIC in the same subjects; and
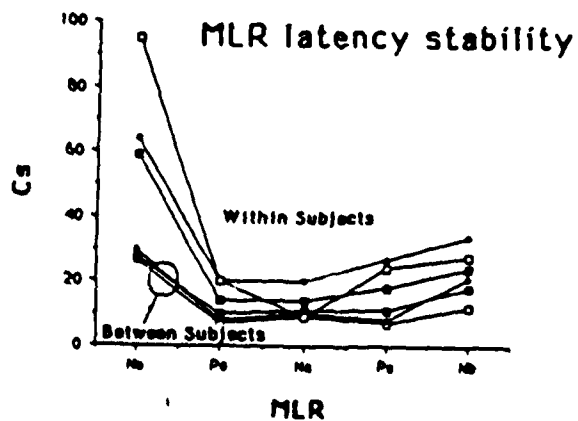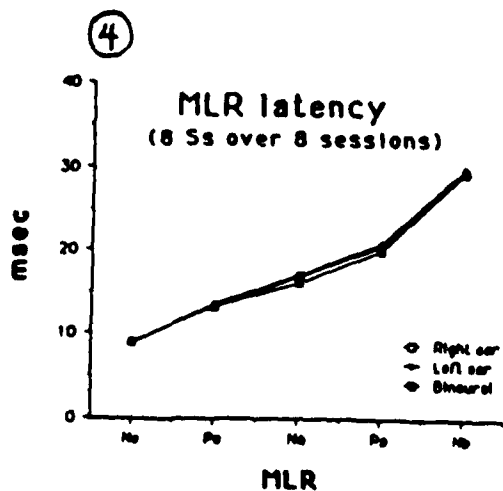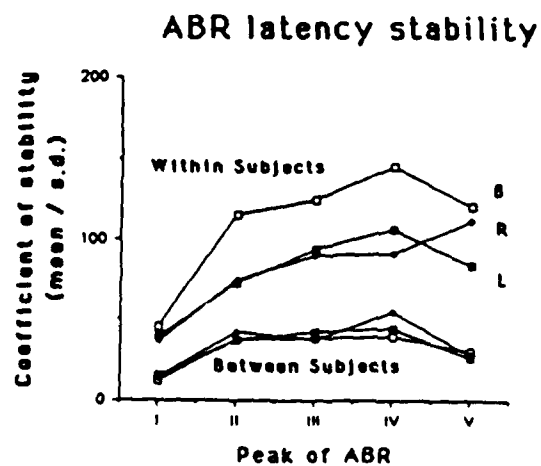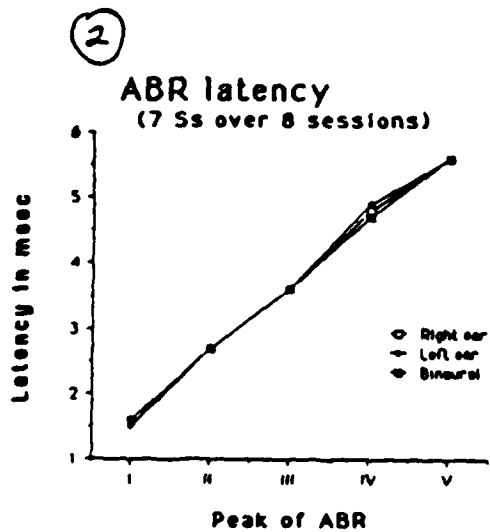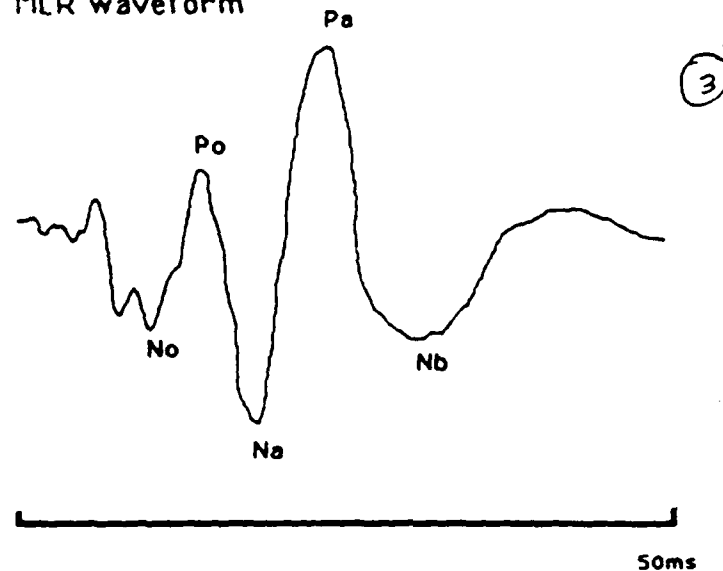
6) test the same subjects with repeated-measures EPs and behaviorally, to determine whether ear differences that can be demonstrated electrophysiologically are reflected in behavioral tests such as dichotic listening.

*Note: The Coefficient of Stability we refer to here is our name for the reciprocal of Pearson's Coefficient of Variation (s.d./mean). In our one published report to date, and our earlier presentations to ASA, we have used the (mean/s.d.) calculation, but referred to this form of the ratio as Coeff. of Var. This practice of using the name of one version of the ratio to also apply to its reciprocal has proven confusing for some audiences. As a result, we here revise our usage: we will retain the (mean/s.d.) form of the ratio, since it emphasizes stability rather than variability, more suitable for the phenomena we are interested in, but will now refer to it as the Coefficient of Stability, to distinguish it from the Pearson version, and to provide a more intuitively descriptive term for an index which increases as the standard deviation decreases.

Published: Lauter JL and RL Loomis (1986) Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability. I. Absolute latencies of brainstem vertex-positive peaks. Scand Audiol 15: 167-172.

① 

MLR waveform

③

Po

Pa

No

Na

Nb

50ms

②

ABR latency
(7 Ss over 8 sessions)



Latency in msec

Peak of ABR

○ Right ear
● Left ear
● Binaural

ABR latency stability



Coefficient of stability (mean / s.d.)

Within Subjects

Between Subjects

Peak of ABR

B

R

L

④

MLR latency
(8 Ss over 8 sessions)



msec

No  Po  Na  Pa  Nb

MLR

○ Right ear
● Left ear
● Binaural

MLR latency stability



Cs

Within Subjects

Between Subjects

No  Po  Na  Pa  Nb

MLR

## ABR and MLR latency stability
(8 Ss over 8 sessions)

⑤

Cs

B

R

L

I    II   III   IV   V    Na   Pa   Na   Pa   Nb

**Peak of EP**

## ABR V and MLR latency stability

⑥

Cs

NN

○ Right ear
+ Left ear
● Binaural

ABR V  MLR Na  Pa  Na  Pa  Nb

**Peak of EP**

Cs

CT

ABR V  MLR Na  Pa  Na  Pa  Nb

**Peak of EP**

## Replication of MLR stability profiles:
## 1st 4 sessions compared with 2nd 4

⑦

Cs

NN
*Right*

2nd 4

1st 4
sessions

Na   Pa   Na   Pa   Nb

**MLR**

CH
*Binaural*

1st 4

2nd 4

Na   Pa   Na   Pa   Nb

WB
*Left*

2nd 4

1st 4

Na   Pa   Na   Pa   Nb
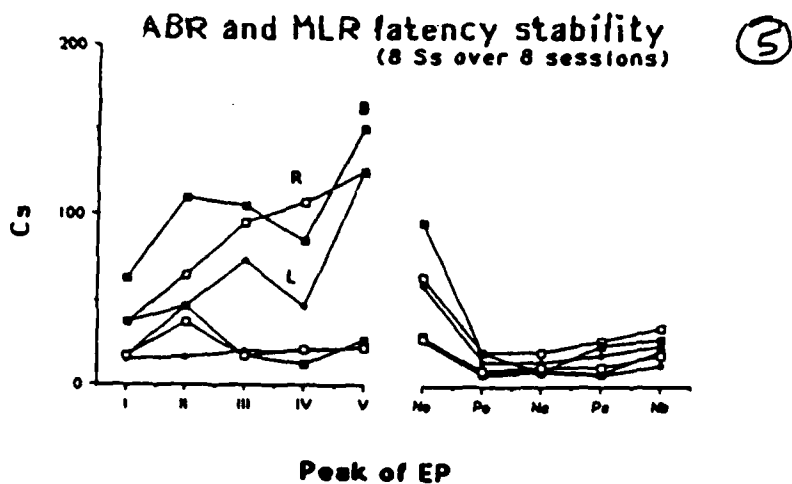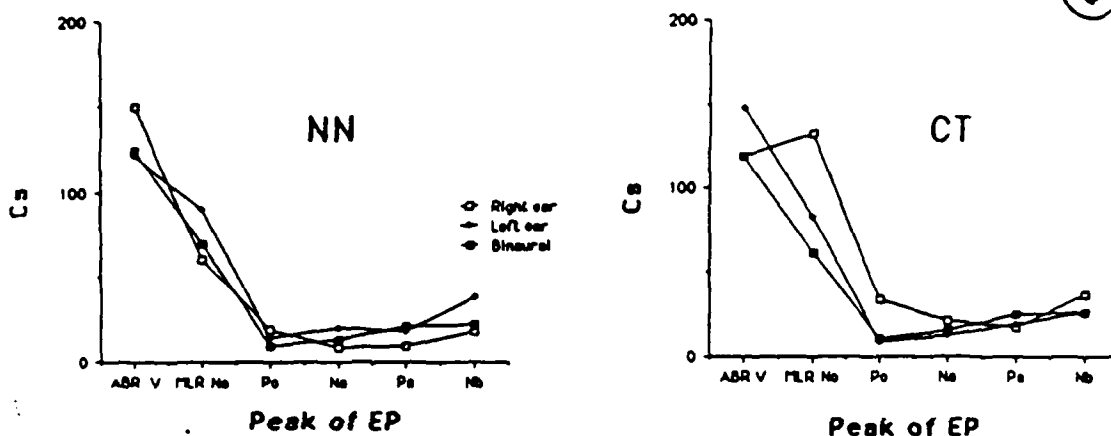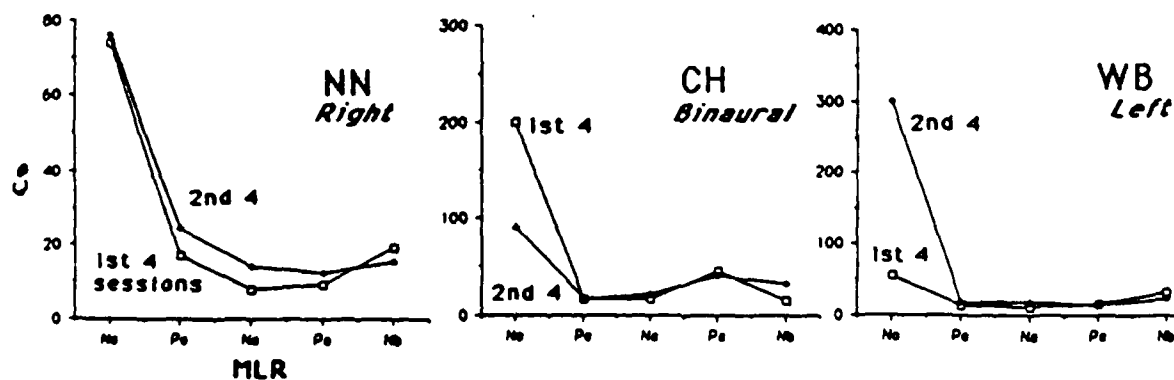
ABR and cortex latency stability
(7 Ss over 8 sessions)
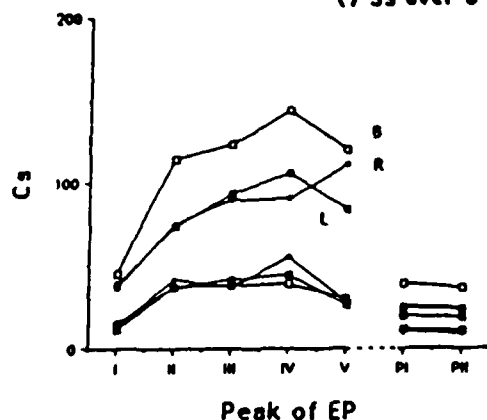


ABR, MLR, cortex latency stability
(ABR, cortex = 7 Ss;
MLR = 8 Ss)
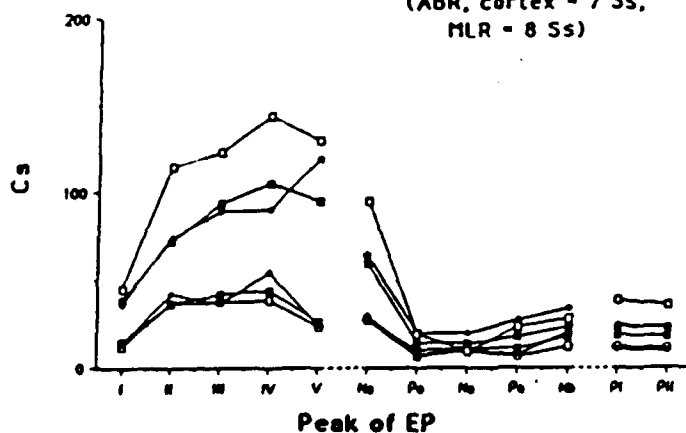
Planned studies using AEP repeated measures:

1. extended time series in the same subjects: 3-4 months

2. compressed time series to collect same size database in less total time (several waveforms per day, several days per week?....)

3. three time windows (ABR, MLR, cortex) measured in the same subjects

4. comparisons of WS and BS stability as a function of age

5. asymmetries in stability profiles compared with Binaural Interaction Component (Berlin et al)

6. comparison of #5 with behavioral tests on same subjects (e.g., dichotic listening)

VOT variability in Mandarin Chinese: interactions with tone. Judith L.
Lauter and Fang Ling Lu, Speech & Hearing Sciences, University of Arizona,
Tucson AZ 85721

Our interests in the sounds of speech as complex sounds have led us to
study the within- and between-subject variability of the acoustical
characteristics of phonemes. We have previously reported measures of
stop-consonant VOT variability in repeated productions of nonsense
disyllables by talkers of American English, Japanese, Mexican Spanish, and
Navajo. Today we will report VOT variability results for six speakers of
Mandarin Chinese, together with data on the characteristics of tone
production in the same tokens.

Six native speakers of Mandarin Chinese, three females and three males,
served as subjects. All were born and raised in Taiwan, and were fluent in
American English.

Subjects were seated in an anechoic chamber, and productions were
digitized using a pulse code modulator and recorded onto the video portion
of video tape.

Each talker was presented with two lists for recording. The first (SLIDE
1) was a list of nonsense disyllables to be read as representing the six
English stop consonants, and a similar second list to be read as Mandarin
stops. The English list was read 6 times. The Mandarin list was then read
once using tone 1 for the second syllable, then using tone 2, etc., for all five
Mandarin tones. Then the entire Mandarin tone-stop set was read through
five more times.

For analysis of VOT durations, recorded productions were played through
a Data Translation board sampling at 20 kHz with 12 bit resolution into a
microcomputer, and stored on hard disk.

Using software developed in our laboratory, the VOT portion of each
token was visualized in a waveform display, and VOT measured as the time
between burst release and the onset of voicing, defined as the point of
transition from a predominance of noise to a predominance of periodicity.

We will first report results of VOT analysis, followed by descriptions of
contour characteristics in both English and Mandarin productions.

The next slide (#2) presents mean VOT values for voiced and voiceless
English stops as produced by the Chinese talkers. These values are very
similar to those used by native English speakers (next slide: #3), as
indicated by this comparison between mean VOT for English stops produced
by the three Chinese females (on the left) and three native American English

female talkers (on the right). Note that both groups of talkers use clearly different VOT values for voiced vs. voiceless stops, but do not use VOT unambiguously to distinguish place.

The next slide (#4) presents mean VOT values for voiced and voiceless Mandarin Chinese stops. These measures are from productions using tone #1, and are typical of VOTs in all tones produced by these talkers. The next slide (#5) compares these VOT values for Mandarin Chinese with those previously reported for female talkers of American English, Navajo, Mexican Spanish, and Japanese. VOT values in Mandarin Chinese are clearly most similar to those of English and Navajo.

VOT variability was described in our previous reports in terms of the Coefficient of Variation, in which standard deviation is considered as a percentage of the mean. The next slide (#6) presents the two VOT variability patterns shown by most subjects studied to date; pattern #1, shown by the four talkers at the top, is characterized by consistent VOT timing in voiceless stops, and relatively higher variability in voiced stops, particularly for /b/; while pattern #2, represented by the bottom two talkers, is characterized by high variability primarily for /t/.

The next slide (#7) presents such VOT variability patterns for the six Chinese talkers, for productions made using Mandarin tone 3. Note that for five of the talkers, the pattern shown is a variation of VOT variability pattern #1 shown on the previous slide: consistent production in voiceless stops, and the greatest relative variability for /b/. Subject YT in the lower right shows a unique pattern. However, these same VOT variability patterns are not used by the subjects for stops produced with all tones. In examining the variability patterns for all tokens, we observed not only differences from tone to tone, but also similarities between the VOT variability in certain tone sets and English.

Specifically, the VOT variability pattern for Mandarin tone 4 and for the English tokens seemed to be similar for several subjects. The next slide (#8) presents a comparison of the VOT variability pattern shown for Mandarin stops spoken with Mandarin tone 4 by each of the female Chinese talkers, with the VOT variability pattern shown by each of these talkers in her English stop productions. For all three talkers, the VOT variability pattern for the English stops is more like that of the Mandarin stops produced with tone 4 than for any other tone.

This similarity between tone-4 VOT variability and English, and the relative dissimilarity between the VOT variability of other tones and English, led us to examine the fundamental-frequency contours used by

these talkers for the Mandarin and the English tokens. We will present only the data for the female talkers here.

Recorded tokens were played through the pulse code modulator into a Kay Visi-Pitch interfaced with an Apple IIe computer and a NEC dot-matrix printer. The next slide (#9) shows examples of contours as reproduced on the computer screen for disyllables representing the five Mandarin tones and English for one talker (MW). Tone 1 is high and flat, tone 2 follows a fall-flat-rise pattern for this talker, tone 3 is falling, tone 4 is a fall-rise pattern, and tone 5 is falling; the English token is produced with a fall-rise pattern similar to that of tone 4. The next slide (#10) shows simplified contours of the 5 Mandarin tones and English based on average values for all six repetitions of the stops by subject MW. These contours were derived by using the Visi-Pitch cursors to determine the initial FO value in each second syllable, the final value, and values at intermediate points where FO changed in direction. Note again the resemblance between the contour used for Mandarin tone 4 and for the English stops.

The next slide (#11) shows similar contours for talker LL. Again, the contour used for the English stops most closely resembles that used for the Mandarin stops produced with tone 4. The next slide (#12) shows contours for talker YT. Although several of her tones resemble the shape of the English contours, i.e., simple falling, more precise measurements related to tone characteristics reveal the close resemblance between tone 4 and English contours. The next slide (#13) presents the combined results of measurements of value of FO at the start of each contour (FO values are indicated with filled symbols), with total duration of each contour (open symbols). Values for tone 4 and English are circled for each talker for ease of comparison. For all three subjects, these characteristics provide further evidence of the similarity between the FO contours used for the tone-4 and English productions.

The last slide (#14) summarizes these observations, comparing for each of the three female talkers, the FO contour for tone-4 and English tokens (on the left) with the VOT variability pattern shown for each set of productions (on the right).

(SLIDE OFF) The relation suggested in these results between stop-consonant VOT variability and characteristics of the fundamental-frequency contour used for the following vowel is not as unexpected as it might first appear. Voice-onset timing in stop consonants is after all contributed to by the larynx, which is also the focus of control for initial settings and planned variations in voice fundamental frequency. An interesting implication of this finding is that it may represent an instance of the

interaction between production activities devoted to segmental aspects of speech, such as VOT timing in stops, and activities related to characteristics such as voice intonation which in some languages contribute to suprasegmental functions of speech such as cueing syntactic patterns and expressing emotion.

Although VOT and tone are both properly segmental for these Chinese talkers, the same interaction may occur in speakers of other languages. Currently we are examining the productions of the male Chinese talkers, to determine if resemblances in Mandarin and English voice contours predict patterns of VOT variability in their stops, and we are planning to examine speakers of English to test whether any systematic variations in voice contour following stops result in changes in VOT variability patterns.

To explore the perceptual implications of these findings, dichotic-listening experiments are planned, testing Chinese and English listeners for identification of both stop consonants and tones. The two types of phonemes are of particular interest for dichotic testing, since they represent two major categories of sound that are clearly distinguished in patterns of relative ear advantages: sounds which change quickly over time and have broad-band spectral characteristics (such as the VOT portion of stops), vs. sounds which change fairly slowly over time and are identified according to narrow-band characteristics (such as voice F0). We expect to find that Chinese listeners, who are familiar with the use of co-varying VOT and F0 patterns as segmental cues in speech, will perform very differently in these tasks compared with English listeners, for whom fast broadband speech sounds and slow narrow-band speech sounds lie on separate sides of the segmental-suprasegmental boundary, and may be perceived in fundamentally different ways.

Presented to International Phonetic Sciences Western Hemisphere Conference, Miami Beach, November 1987.

Fig. 1

'5

**English stop consonants:**

| | |
|---|---|
| ha-ba | (as in "Bob") |
| ha-pa | (as in "pop") |
| ha-da | (as in "doll") |
| ha-ta | (as in "Tom") |
| ha-ga | (as in "God") |
| ha-ka | (as in "cod") |

Fig. 2

## VOT duration: English stops
### (6 Mandarin Chinese)
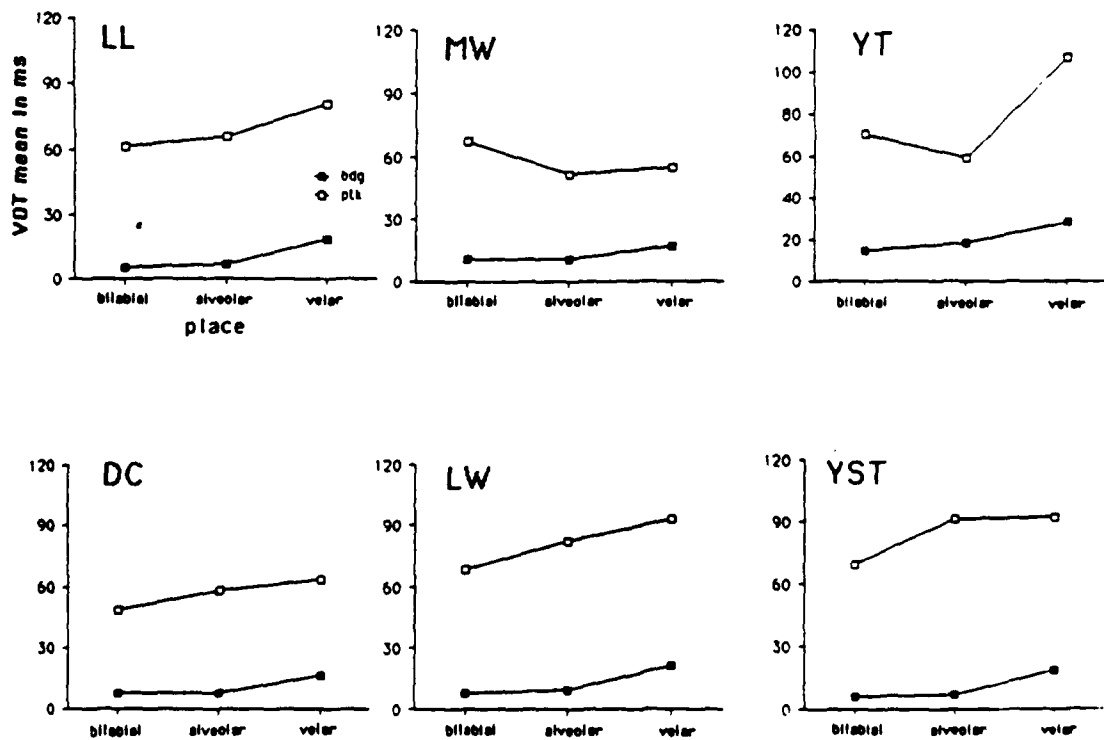
Fig.3

VOT duration: English stops
(3 Mandarin vs. 3 American English)
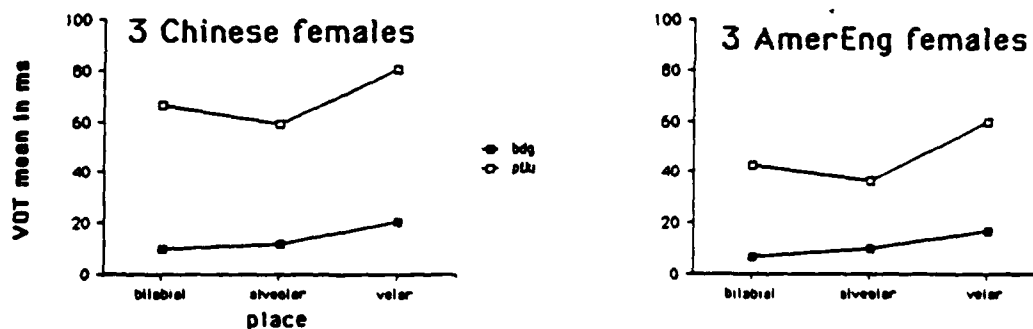
Fig.4

VOT duration: Mandarin 1
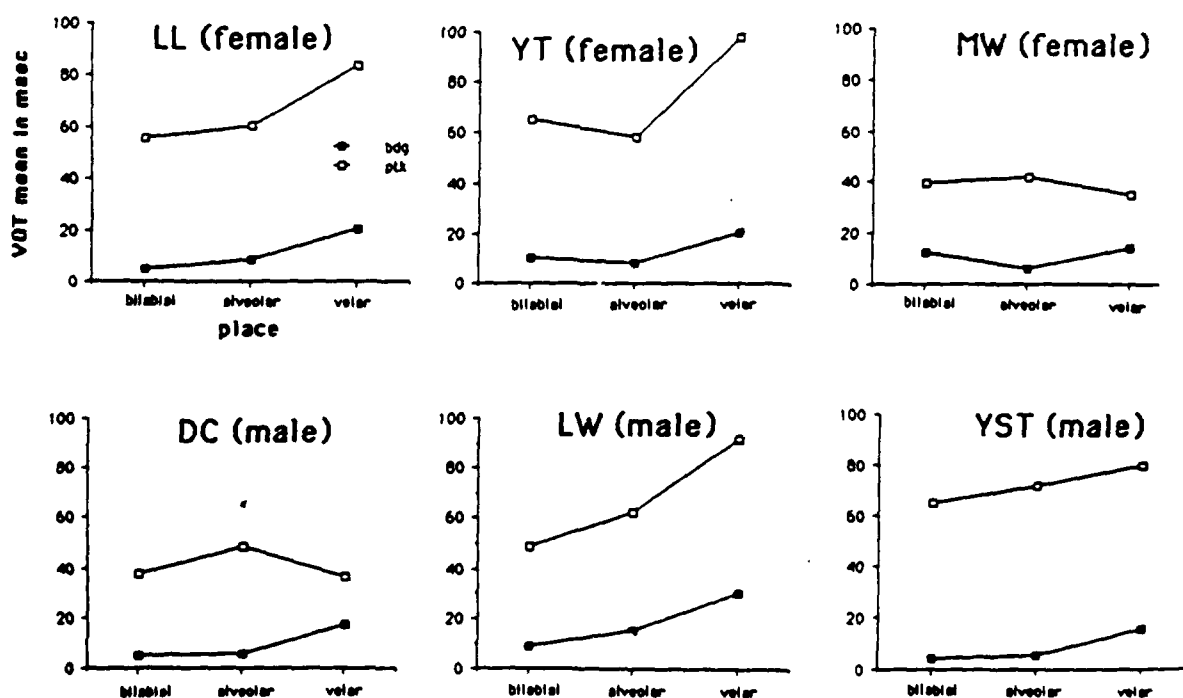
Fig. 6

## VOT variability (V)
(examples from 4 languages)

AmerEng: CB

Navajo: JJ

AmerEng: NP

AmerEng: JL

MexSpan: JW

Japanese: YW

place

VOT V (s.d./X)×100

Fig. 5

## VOT duration: 5 languages
(3 females each)

Navajo

American English

Japanese

Mandarin (tone 1)

Mexican Spanish

place
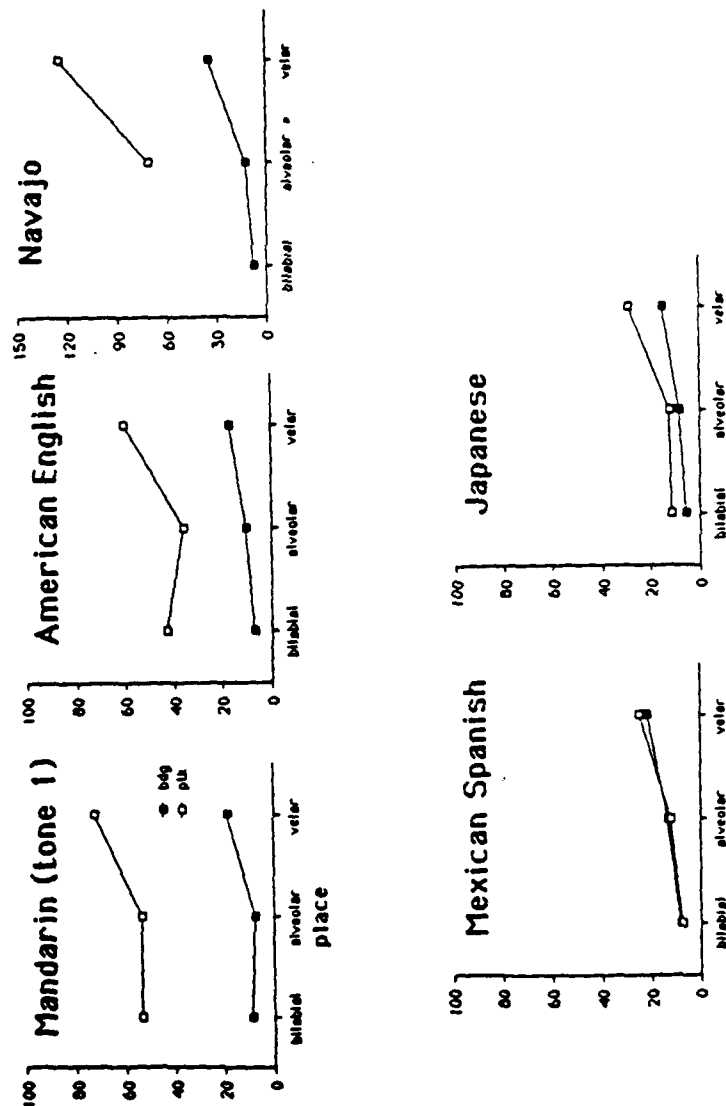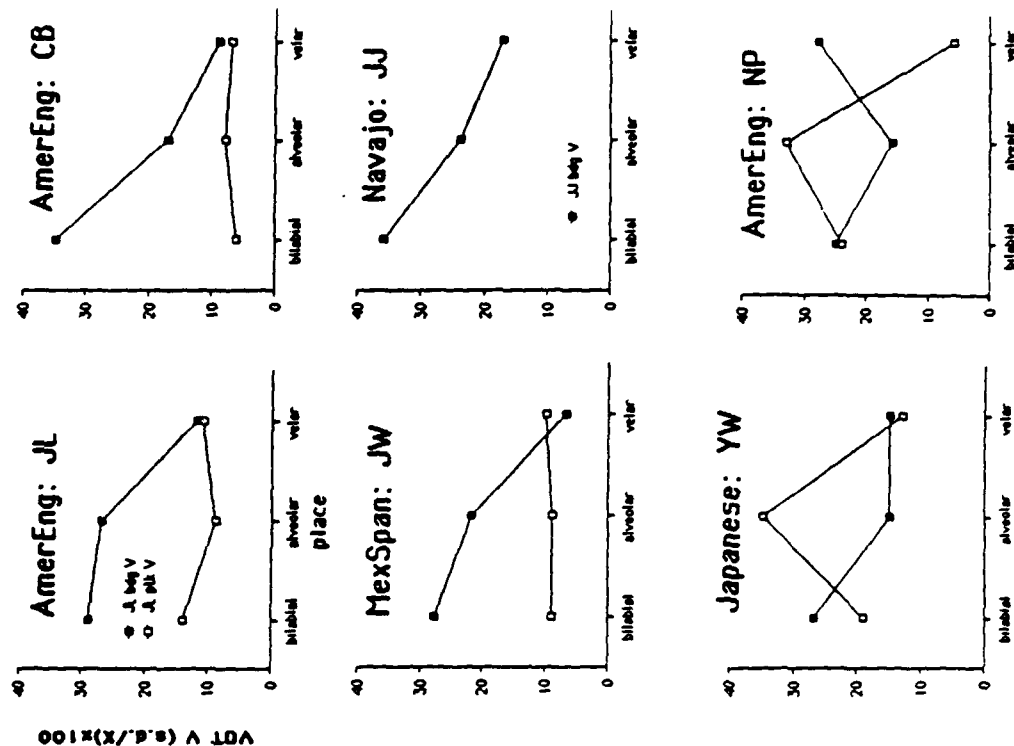
Fig. 7

# VOT variability (V)
## (Mandarin 3)
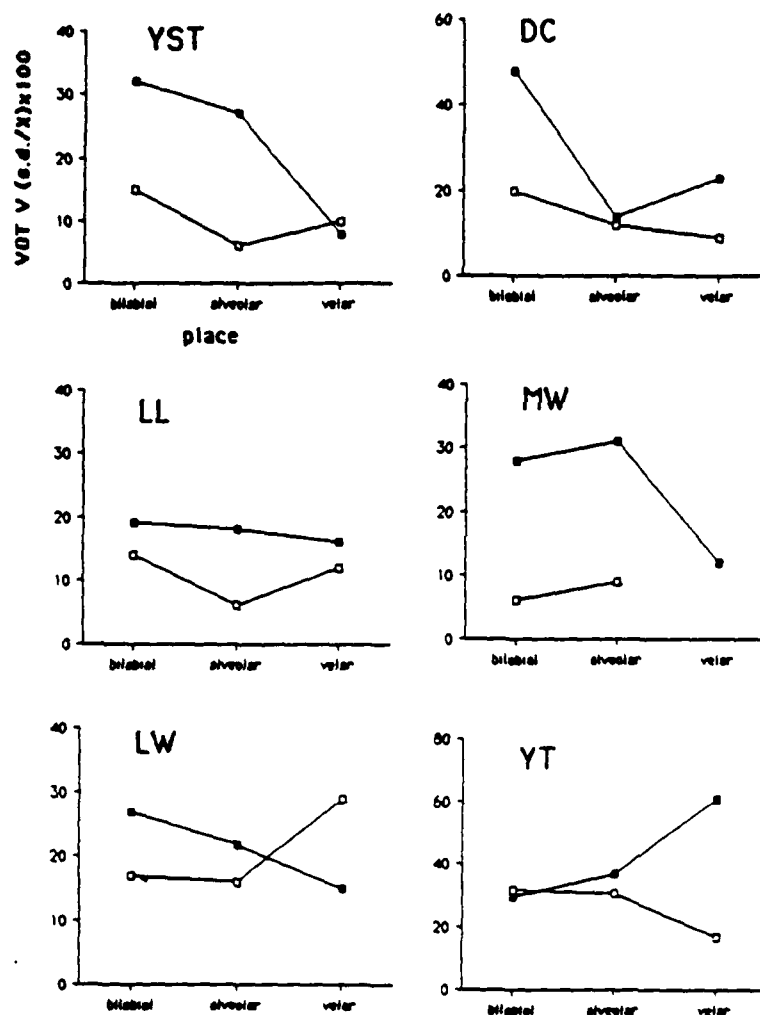


Fig. 8

# VOT variability (V)
## (English vs. Mandarin 4)
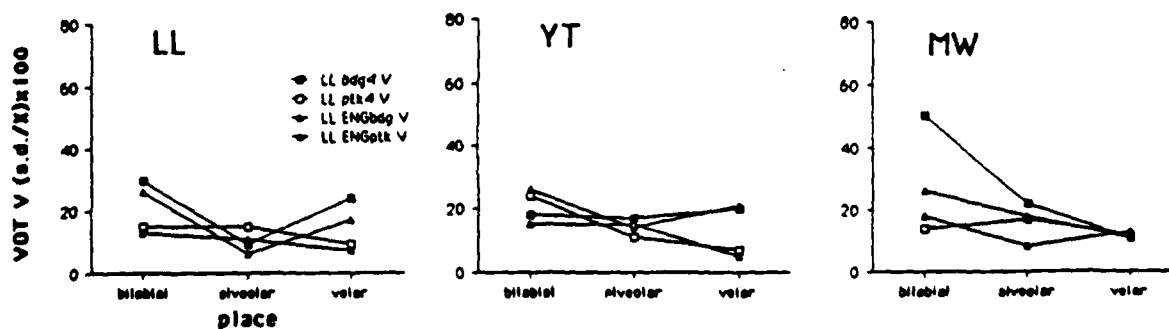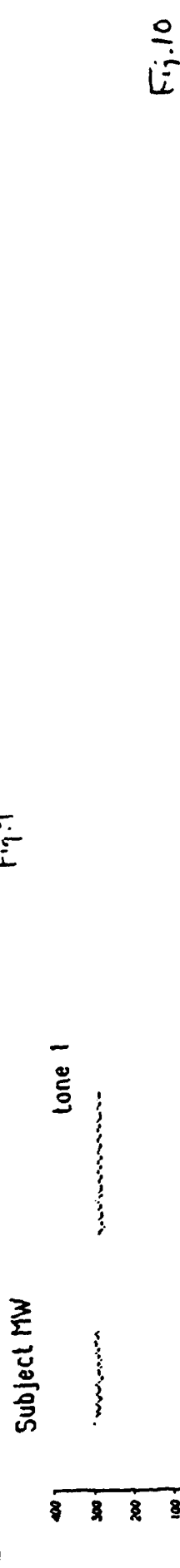
Fig. 10

Fig. 9

FO contour: simplified shape
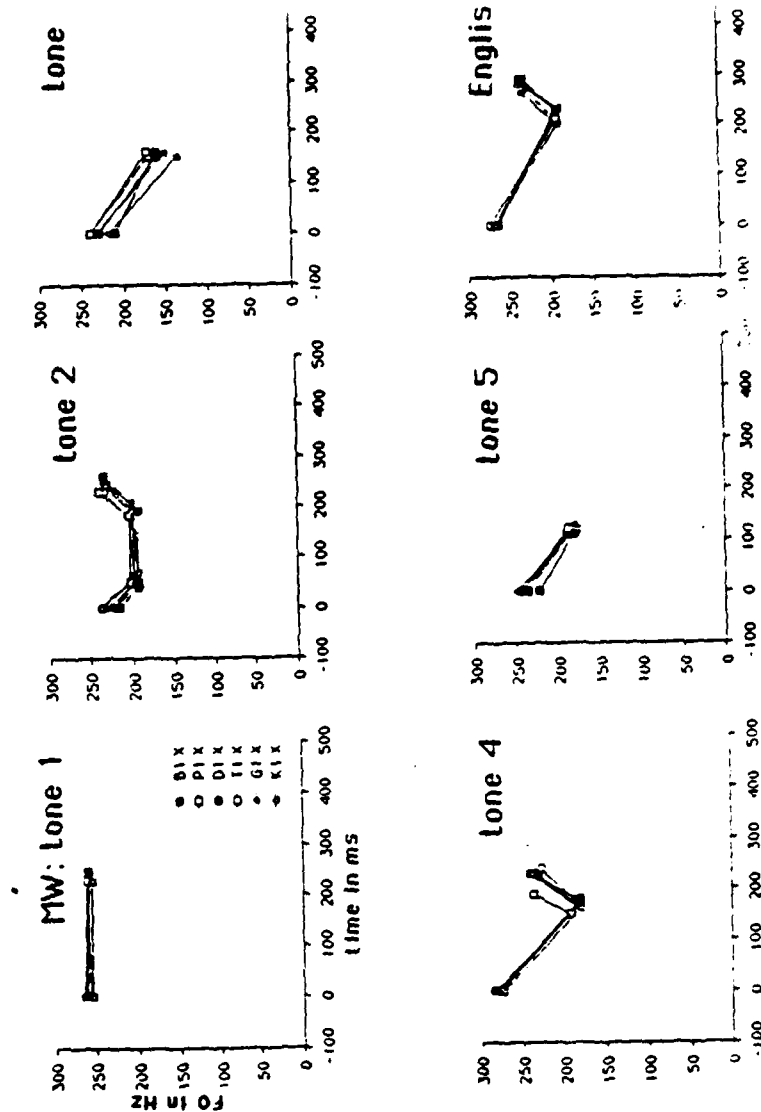
# FO contour: simplified shape
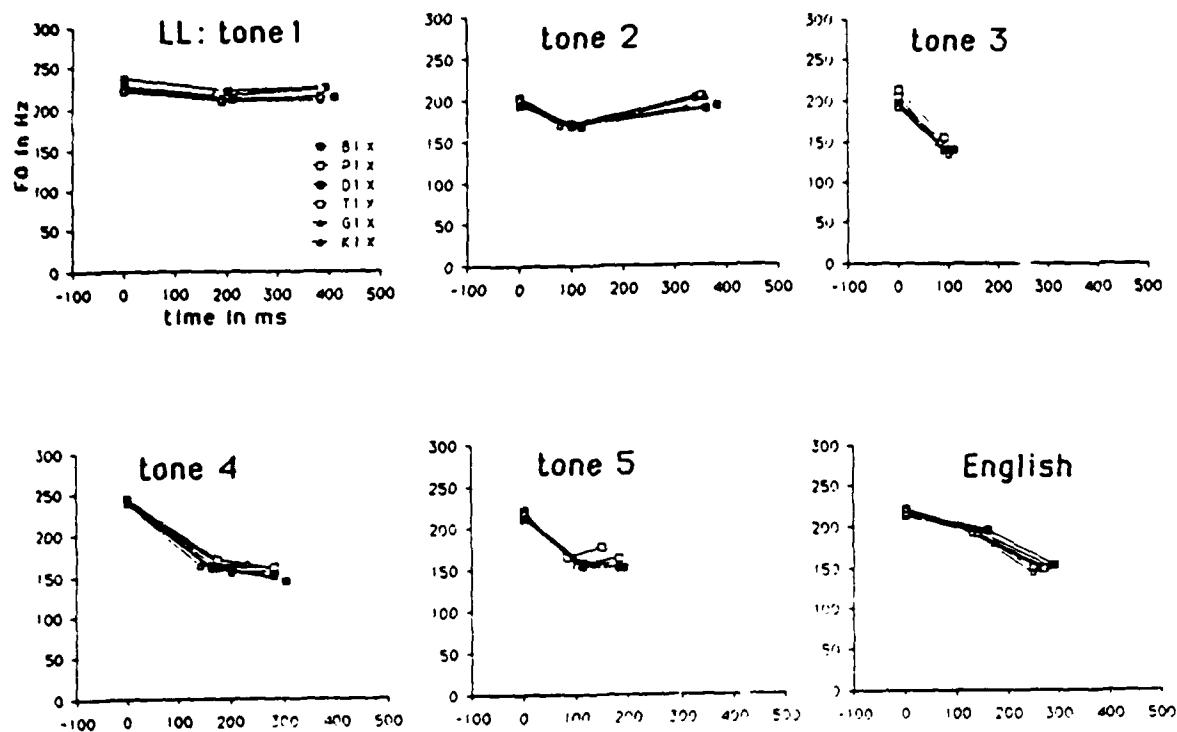
Fig. 1



# FO contour: simplified shape
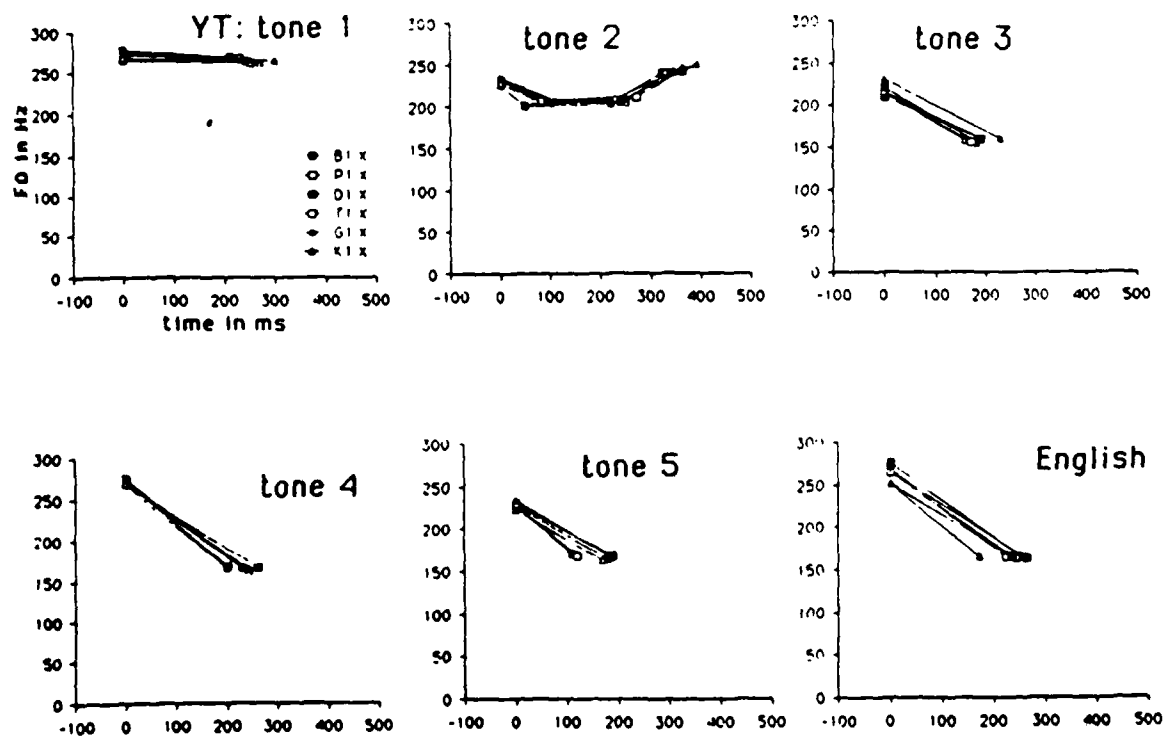
Fig. 1

Fig. 14
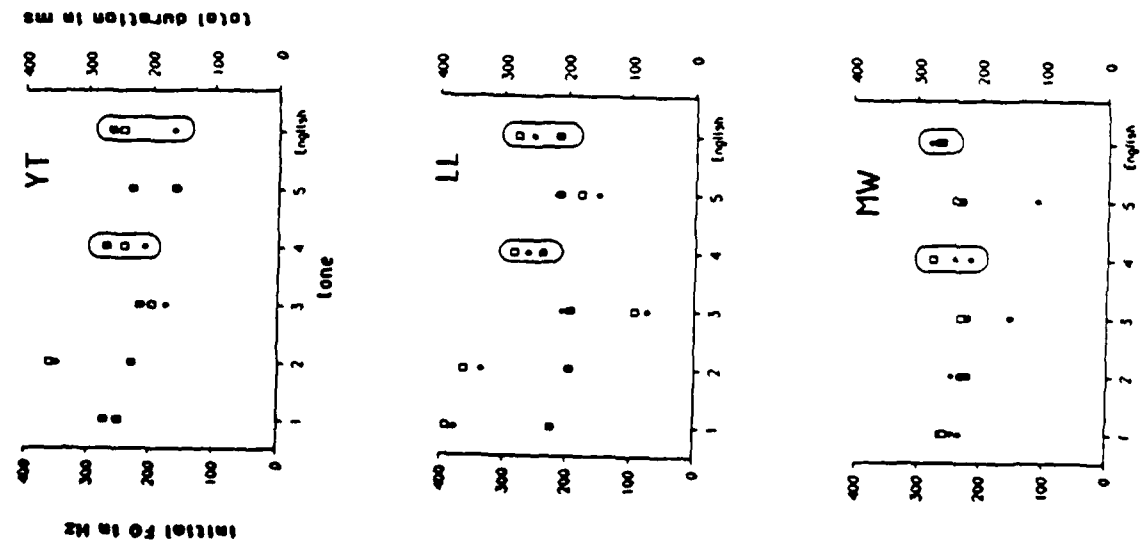
FO contour and VOT variability

(English vs. Mandarin 4)
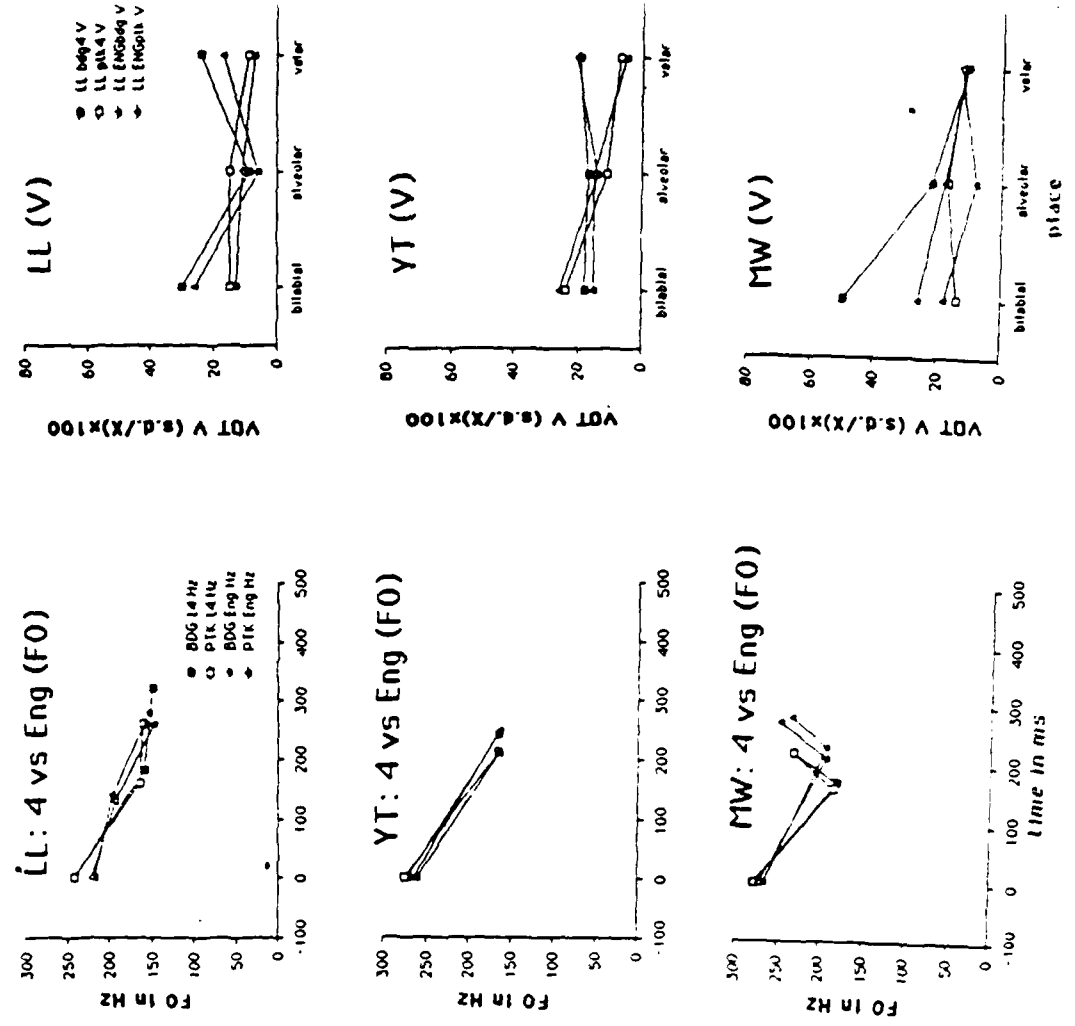


Fig. 13

FO contour: initial frequency
and total duration

(5 Mandarin tones + English)

**Windows to the brain:**
**what contemporary imaging devices can**
**reveal about speech and hearing**

**Judith L. Lauter, Ph.D.**
**University of Arizona**
**Tucson AZ 85721**

**A WORKSHOP**
**Sponsored by:**
**Department of Speech & Hearing Sciences**
**University of Arizona**
**Tucson AZ**

**February 18, 1989**

WORKSHOP HANDBOOK

COURSE DESCRIPTION

Over the past five to ten years a number of devices for noninvasive, high-resolution study of human brain anatomy and physiology have been developed. The great potential of these devices for providing new insights into human function including communication and communication disorders can be difficult to appreciate because descriptions of the devices often present a bewildering array of technical details and new vocabulary. This workshop will attempt to render the new techniques more "user friendly," by providing a brief introduction to the design and methods associated with several new approaches, illustrating with actual photographs—in color where available— the "window" on brain function that each provides, and discussing the ways in which these new devices can help us understand more about human communication.

Objectives:

1. Introduce the vocabulary, basic methodology, and principles of testing associated with: Magnetic Resonance Imaging (MRI), quantitative EEG brain mapping ( qEEG), new uses of evoked potentials (e.g., for documenting individual differences, including right/left asymmetries), Magnetoencephalography (MEG), and Positron Emission Tomography (PET).

2. Illustrate through the results of actual experiments, including photographs of sample images, the type of information relevant to human speech, language, and hearing, that each technique makes available, and how these types of information are related.

3. Emphasize the need for an attitude of healthy scepticism combined with cautious optimism regarding these new techniques: what questions can be asked with them, which cannot, how to tell the difference.

4. Look toward the future, considering new developments that may be expected, and discussing the role that speech-language-hearing professionals can play in guiding the application of noninvasive devices to the study of human communication and its disorders.

INTRODUCTION

Several recent technological and methodological developments offer students of human behavior an unprecedented opportunity to look inside the "black box" of the living brain. These technologies include: 1) high-resolution anatomical scanning provided by Magnetic Resonance Imaging (MRI); 2) quantitative analysis of EEG activity (qEEG) with topographic mapping based on computer-mediated spectral analysis of multi-channel scalp-electrode recordings; 3) individual subject characterization in terms of evoked potentials (EPs) based on repeated-measures testing, and waveform combination and derivation; 4) detection of magnetic-field generation, the result of electrical activity in discrete brain regions, via Magnetoencephalography (MEG); and 5) three-dimensional autoradiographic maps of the cranial central nervous system based on methods such as Positron Emission Tomography (PET).

In this workshop, we will first introduce each of these methods in terms of vocabulary, techniques, and procedures involved in preparing and testing subjects. Then data will be presented illustrating how each approach might be useful for studying aspects of brain function related to human communication. Although the focus in the reported results will be on hearing, we will also review findings for visual and motor stimulation where available. In the subsequent discussion, we will provide a bridge between the results on normal subjects reported in our review, and possibilities for clinical applications, considering the realities of demands on subjects, as well as the appropriateness of the picture of the brain provided by each method for studying different types of disorders.

Outline of the Workshop:

(morning session)
Dr. Lauter: Introduction to noninvasive imaging methods
MRI and electrophysiological methods
PET results
(afternoon session)
Dr. Lauter: The CNS Demonstration Project
Panel: Discussion of applications to communication disorders
and audience questions and comments

Table I: Comparisons of methods

| Name | brain feature | dimensions | S interface | time/pic | time/sess |
|------|--------------|-----------|-------------|----------|-----------|
| MRI | anatomy | 3-D | cylinder(ext) | 10 min | 60 min. |
| EPs | electricity | 2-D | electrodes(on) | 2 min. | 60 min. |
| qEEG | electricity | 2-D | electrodes(on) | 5 min. | 60 min. |
| MEG | magnetic fields | 2-D | dewar(ext) | 5 min. | 60 min. |
| PET | metabolism | 3-D | isotope(int) | 2 min. | 90 min. |

Figure 1. Comparisons of methods based on spatial resolution, and temporal resolution (minimum time per scan).

## Spatial resolution scale

x 15 microns

$10^6$

$10^5$ > behavioral methods

$10^4$ > qEEG, EPs

$10^3$ > PET

$10^2$ > MEG, MRI/MRS

$10^1$ > cortical columns

$10^0$ > single cell

## Temporal resolution scale



device

be — behavioral
mri — MRI
meg — MEG
ep — EPs
eeg — qEEG
pet — PET

$10^0$   $10^1$   $10^2$   $10^3$

time in min.
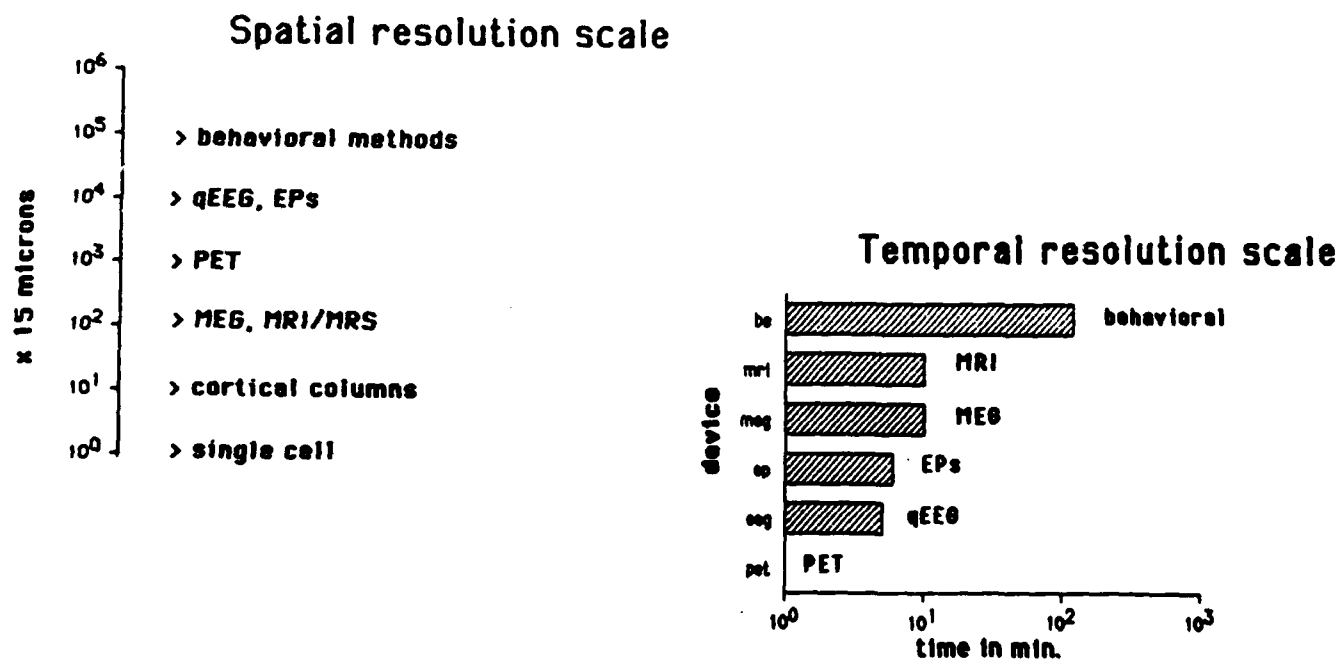
The two resolution scales suggest that there are trade-offs in these two types of resolution that make some devices more useful than others. For example, although magnetic resonance spectroscopy (MRS: a physiological version of MRI) offers excellent spatial resolution (approximately 1 mm), the time required for data collection (10-20 minutes) seems daunting from the point of view of neurophysiological applications. On the other hand, although the PET scan is an order of magnitude more coarse in terms of spatial resolution (approximately 1.5 cm), its temporal resolution (using oxygen-15, 40 sec) makes it very attractive for studies of normal brain function.

Until more testing is done, there is no way to judge which device, if used alone, would provide the most informative results for students of human brain function. The view of the brain collected with "poor" temporal resolution (e.g., MRI/MRS) may in fact be more appropriate for comparisons with behavioral results that are collected over a similar time scale. Similarly, results collected in "coarse" spatial resolution (e.g., PET's 1.5 cm) may reflect a level of brain organization that is more relevant to understanding behavior than the "microneurophysiology" provided by microelectrodes.

MAGNETIC RESONANCE IMAGING (MRI)

The application of Nuclear Magnetic Resonance (NMR) physics to anatomical imaging has had a revolutionary effect on modern medicine and promises unsuspected insights into the structure and function of the normal brain (cf. Brant-Zawadzki & Norman 1987).

The subset of principles underlying MRI technology include: 1) an element's randomly spinning protons tend to align with a superimposed magnetic field; 2) different compounds containing that element are characterized by different "magnetization" times (also called "spin-lattice relaxation"), or T1, which equals the amount of time required for 63% of the relevant protons to align; 3) a radiofrequency signal (RF) can affect the pattern of this alignment such that all protons are put into a "high-energy state;" 4) at cessation of the RF excitation signal, the protons relax (at different rates) to their pre-RF (but still magnetized) state, and in doing so, emit an RF signal ("spin echo"), which can be detected and utilized to generate maps of the distribution of spin-echo values within the field of view, in terms of gray-scale-coded images; 5) the intensity of the spin echoes depends on a number of variables (T1 of the different materials involved, amount of time allowed for magnetization, time in the decay envelope where sampling is done), such that data-collection procedures can be used to selectively manipulate contrasts in the resulting images (e.g., gray matter as darker, white matter as lighter).

Subject preparation is minimal, consisting primarily of ensuring that no ferrous metals are present on or within the subject that might be attracted by the strong magnetic field generated by the MRI machine. The subject reclines with eyes closed on a comfortable table with head and upper torso inside the cylinder of the machine. Testing is done in the machine room with lights lowered, but with substantial ambient noise resulting from the RF pulses. The experimenter sits at an operator's console outside the test room. Testing can be done by a single person, but two are preferable, one to monitor the subject and one to operate the equipment. One set of images for the whole brain can be collected in 10 min. Analysis can be done using the data-collection device's standard routines, or on a satellite microcomputer-based system fitted with commercially-available video-capture hardware and image-analysis software. MRI anatomical testing has been accepted as a clinical extension of X-ray and CT scanning, and is currently third-party reimbursable.

Figure 1. Copy of dissection-quality anatomical images taken with MRI.
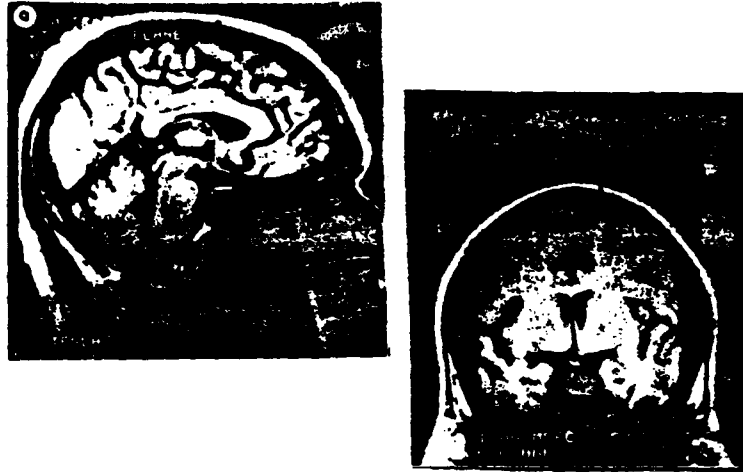


Recently, MRI has been used to examine living brains for details regarding anatomical asymmetries such as have been demonstrated previously using autopsy material (Witelson & Pallie 1973, Wada et al 1975, Geschwind & Levitsky 1968, Teszner et al 1972Rubens et al 1976, Galaburda et al 1978) or CT scans (e.g., LeMay 1977, Chui & Damasio 1980, Pienladz & Naeser 1984)--cf. Fig. 2.

Figure 2. Example of anatomical asymmetries studied using autopsy brains (Geschwind & Levitsky 1968).
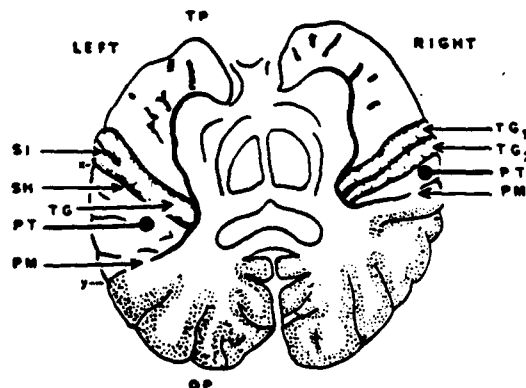
MRI demonstrations of anatomical asymmetries have been directed not only to regions of cortex (Kertesz et al 1986), but also to subcortical centers that may be of importance in language acquisition (e.g., Courchesne et al 1987, Courchesne in press). A number of these authors have speculated that the observed anatomical asymmetries may be related to functional asymmetries such as those that can be observed in behavioral tests, such as dichotic listening and visual-half-field testing, or which express themselves in the sequelae of brain damage.

Attempts have been made to examine the degree of correlation between behavioral and anatomical measures of asymmetries. Ratcliff et al (1980) reported that there were correlations between the angle of the posterior branch of the middle cerebral artery and language dominance as tested with carotid injection of sodium amobarbital; Witelson (1982) tested subjects pre- and post-morbidity and found that for most of her right-handed subjects, a larger left planum temporale was correlated with behavioral right-dominance on a set of behavioral measures (EAs, hand preference, finger tapping); and Kertesz et al (1986) found that although their EA, VHF, and dot-tapping test results were not correlated with MRI-determined asymmetries in the size of the planum temporale, hand preference (rated via five questions regarding common hand usage) was--i.e., right-handers had a larger left planum temporale, while this area in left-handers was the same size on both sides.

A method related to MRI, but relevant to physiology rather than anatomy, is Magnetic Resonance Spectroscopy (MRS). MRS procedures which might prove useful for studying brain function are only just being developed for use in human subjects. MRS holds promise for studying issues related to speech and hearing because it is noninvasive, requires minimal subject preparation, and combines the excellent spatial resolution of MRI with fair temporal resolution (1 scan in 10 min.). MRS is based on observing changes in the chemical spectrum of body tissues, and has been employed to study several organs in living humans, including the brain (cf. Valk et al 1985 and Cohen 1987). However, to date there is no published report describing the use of MRS to examine changes in chemistry associated with functional activation such as responses to sounds. A preliminary experiment conducted at the Waisman Center using new MRS test procedures, including pre-imaging with MRI [Lauter, unpublished report] suggests that MRS has great potential as yet another noninvasive method for studying human brain function.

EVOKED POTENTIALS (EPs)

Research and clinical use of EPs have depended on the development and availability of computers, since the form of EPs most often used represent averages of brain electrical activity time-locked to a stimulus (for a historical review, cf. Davis 1976). In the auditory system, EPs are studied in terms of three response-averaging windows: 1) stimulus onset up to 10 msec post-stimulus-onset, presumed to reflect processing in auditory brainstem nuclei (the auditory brainstem response, ABR); 2) from 10 to 100 ms post-stimulus-onset, involving timing appropriate for responses in auditory nuclei from upper brainstem into primary cortex (middle-latency response, MLR); and 3) beyond 100 msec post-stimulus-onset, where the major contribution is assumed to be from cortical neurons (cortical response, often called auditory evoked potential, AEP, to distinguish the cortical auditory response from cortical somatosensory, SEP, or cortical visual, VEP).

Procedures include placing electrodes on the subject's head, either as a small set of individual electrodes (e.g., at vertex, forehead, both earlobes), or via an electrode cap such as that used for EEG (see next section). Subject preparation requires 15-30 min.; and one response (e.g., to a right-ear click at one intensity level) can be collected in 2 min. Testing is done with the subject reclining in a quiet, darkened room, with the experimenter and test equipment outside. Preparation and testing can be accomplished by one person, though two are preferable, with one to monitor the subject and one to conduct the tests. Analysis can be done while waveforms are still in computer memory, or responses may be stored on disk for offline analysis. There are some commercially-available microcomputer programs which would allow analysis on a satellite machine. EP testing is third-party reimbursable.

Of the auditory responses, the ABR has proven to be most useful in speech and hearing applications, due to its resistance to changes in subject state, such as cooperativeness or state of arousal. In addition to the information available from ABRs using conventional test procedures, further details regarding the characteristics of individual systems are obtainable using new approaches, e.g., variability measurements (e.g., Lauter & Loomis 1985; in press) and waveform comparison (e.g., Berlin et al 1984). Both of these techniques can be used to reveal differences in responses to left-ear vs. right-ear vs. binaural stimuli, and may be helpful for assessing central auditory dysfunction. [Note: some of the new computer-based EEG machines described in the next section can generate topographic color-coded maps from EP recordings, but this workshop will not consider this type of EP result.]

Figure 1. Individual differences ("between-subject variability") and within-subject consistency of ABR waveforms (Lauter & Loomis 1985).

**ABR waveforms**

7 subjects x 3 sessions

binaural 100μsec clicks

11.1/sec 60 dB nHL
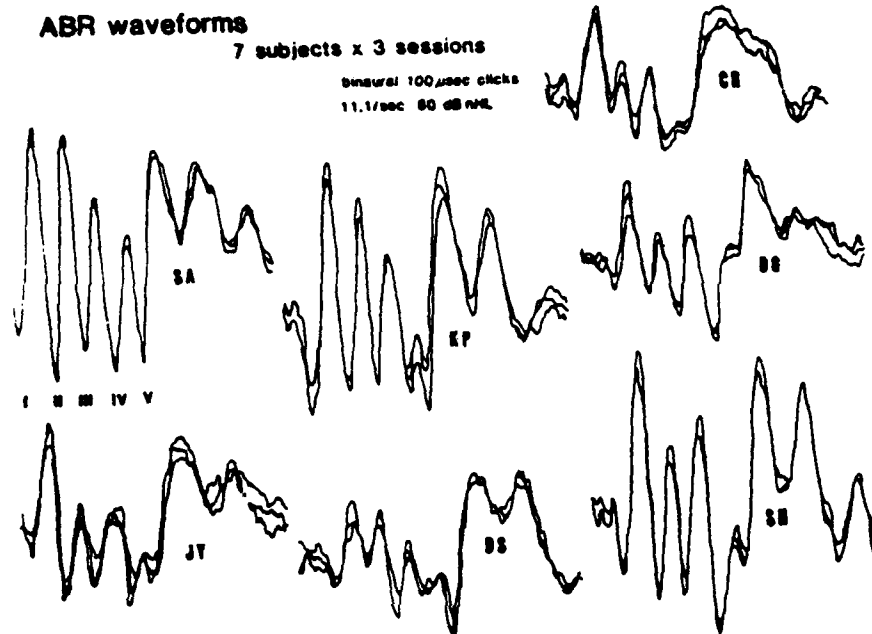
Figure 2. Comparison of absolute values of ABR peak latency (shown on left) and measures of peak-latency variability (on right), with regard to sensitivity to the "three auditory nervous systems:" right, left, and binaural. (Data averaged for eight normal subjects; taken from Lauter & Loomis series.)

**ABR latency**

(7 Ss over 8 sessions)

Right ear
Left ear
Binaural

**ABR latency stability**

Within Subjects

Between Subjects
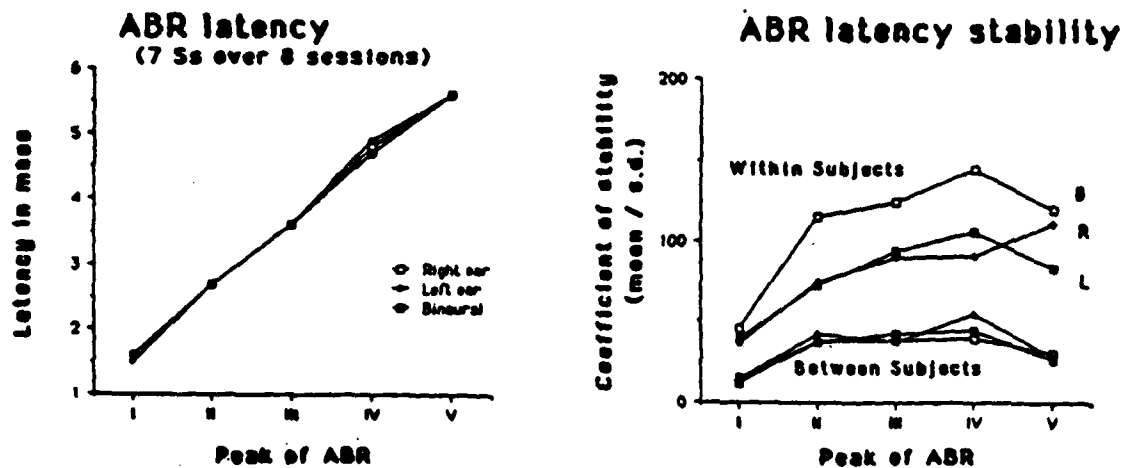
Figure 3. "Variability profiles" such as those shown in Fig. 2, compared for ABR, MLR, and cortex. [Data from Lauter & Loomis series.]
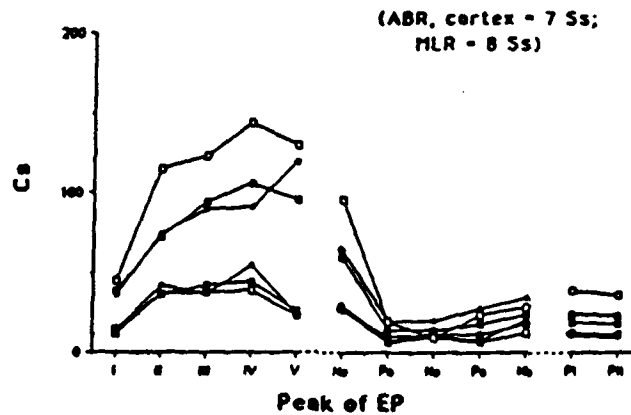


Figure 4. Individual "variability profiles" showing the degree of replicability: compared are profiles calculated for an initial set of 4 sessions vs. profiles calculated for a second 4 sessions [data from Lauter & Loomis series]. Note that in some cases, a subject's profiles for both sets of sessions are very similar (e.g., for latency, SH binaural), while sometimes a subject's profile requires the second set of sessions to resemble that shown by another (e.g., for amplitude, KP's Binaural-Right profile in the 2nd set of sessions looks like the Binaural-Right profile shown by SA in both sets).

Panel A: latency profiles          Panel B: amplitude profiles
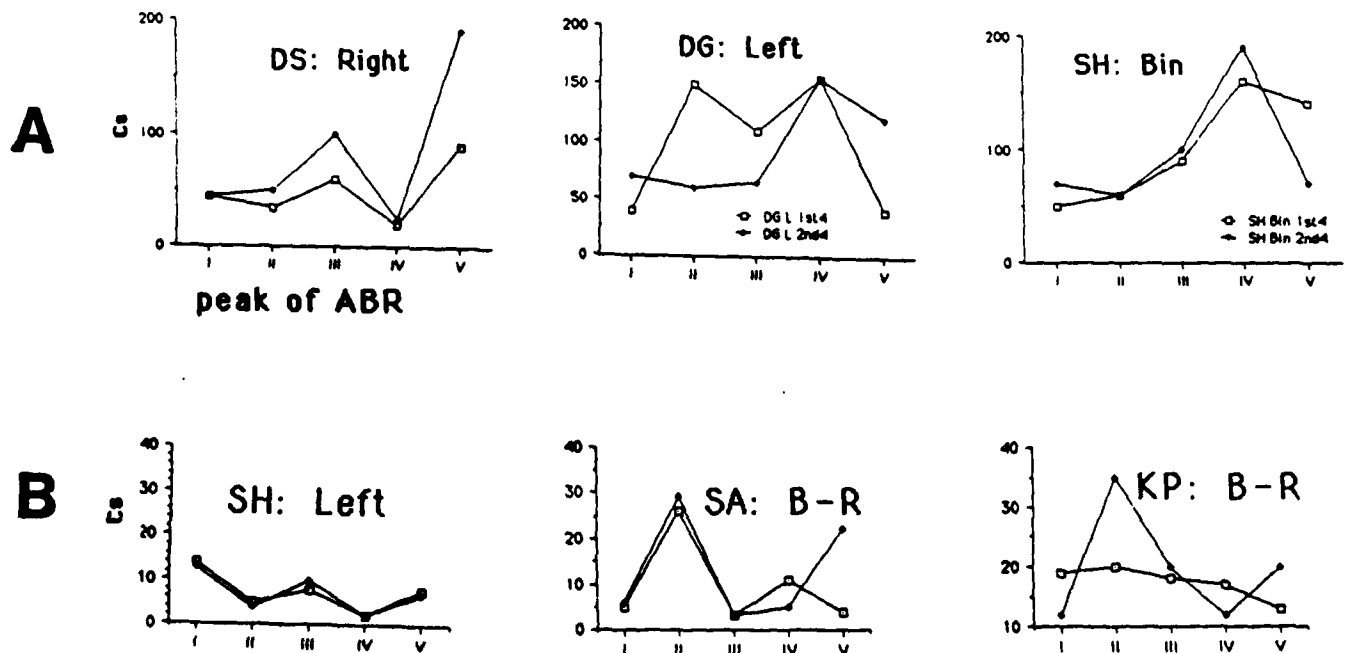
Figure 5. Measures of auditory asymmetries based on ABRs: Panel A: amplitude asymmetry at ABR Peak III based on variability measures (Lauter & Loomis, in press)--cf. amplitude differences at Peak III reported by Levine & McGaffigan 1983); Panel B: "binaural interaction component" at 6 msec post-stimulus-onset measured using waveform addition and subtraction methods described by Berlin et al (1984).
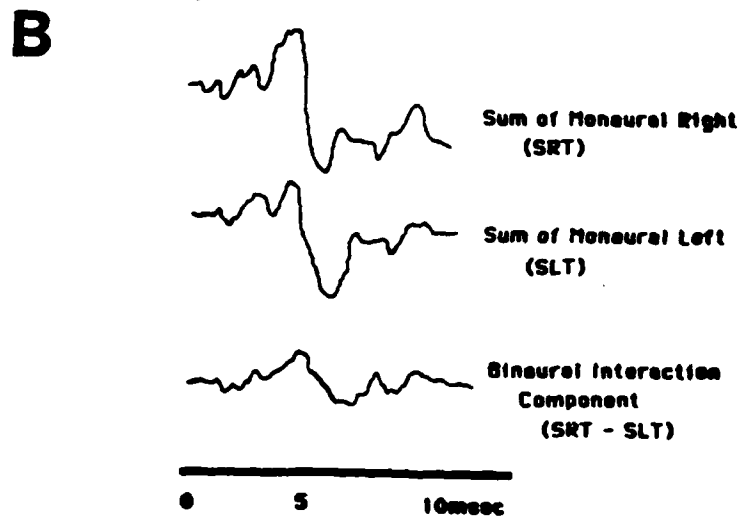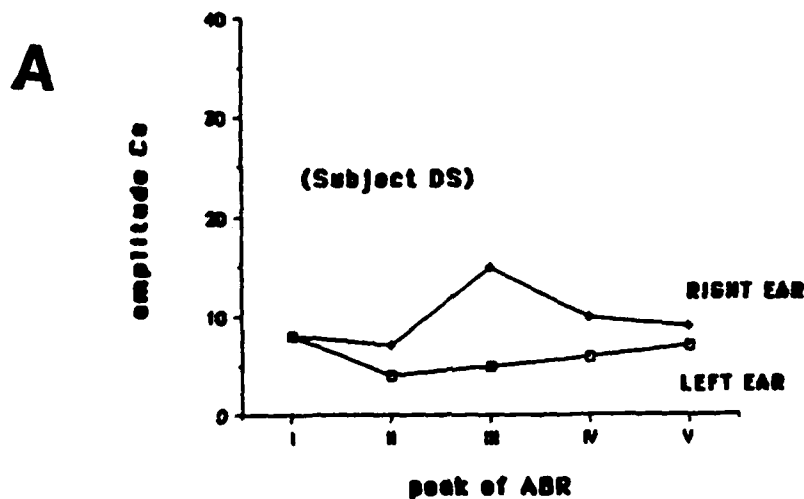
QUANTITATIVE ELECTROENCEPHALOGRAPHY (qEEG)

The oldest method for monitoring human brain physiology is electroencephalography (EEG), accomplished using electrodes placed on the scalp (for a history, cf. Brazier 1977). In addition, spectral analysis of EEG waveforms recorded at specific electrode locations has been utilized to study both normal and abnormal brains. However, new computer-based EEG machines called "brain mappers" provide greater ease of EEG collection and analysis. These machines not only streamline EEG measurement procedures, but also provide color-coded topographic-maps showing simultaneous activity at many locations.

Procedures include fitting the subject with an electrode cap, which ensures standardization and ease of electrode placement. Electrodes are also placed on the earlobes as references, and around the eyes for monitoring eye movements, the artifact most disrupting to EEG. Subject preparation requires 20-30 min., and data collection requires as little as 5 minutes per condition in cooperative subjects. Longer scan times are required when there is much muscle activity, to ensure acquiring at least 1.5 minutes of artifact-free EEG. Testing is done with the subject reclining in a quiet, darkened room, and the EEG machine and experimenter located outside. Preparation and testing can be done by a single person, though two are preferable. Analysis requires the data-collection device; there is only one system currently which provides for satellite analysis, via vendor-supplied software designed to run on a compatible microcomputer system. New EEG methods are considered as extensions of older ones, and are third-party reimbursable.

The approach provided by the new computer-based systems is referred to by several names: neurometrics (John 1977), brain electrical activity mapping (BEAM: Duffy et al 1979), spectral EEG, quantitative EEG, computer-aided analysis of EEG, brain mapping, and EEG topography. Like conventional EEG, the promise of this approach is that EEG can be recorded under both resting and activation conditions, and then compared. Although such procedures have been used in the past in attempts to monitor "cognitive processing," these experiments have been disappointing (cf. Gevins and Schaffer 1980). Studies are just beginning which are designed to explore the uses of qEEG for studying brain function related to sensory and motor processing, with promising applications for speech and hearing.

Figure 1. Left-hemisphere electrode placements according to the "10-20" system
for EEG recordings. Right-hemisphere homologues have even numbers, e.g.,FP2, T4.
Figs. 2 - 4 will present data for electrode locations T3/4 (auditory) and F7/8
(motor). For comparison, Inset A shows the Brodmann map, where primary-
auditory areas 41 and 42 may underlie electrode T3, and motor areas of the frontal
lobe may underlie electrode F7. Inset B shows the human motor-cortex
"homunculus"--note location of the hand area, for looking at our recordings at
electrode position F7 during tests of hand movement.



A



B

Figure 2. Typical table of numbers from a "brain mapper" machine (Cadwell Spectrum 32), analyzing 2.5 minutes of EEG collected during a condition in which the subject was listening to dichotic melodies and mentally labelling those in the left ear. The table indicates the wealth of information available with these computer-based devices. Note that useful interpretation of qEEG results can be based on only a few of these numbers—figures 3 and 4 were compiled by taking only one number, (power asymmetry between electrodes T3 and T4) from this type of table, and comparing that number for different conditions tested in the same subject. [From study reported by Lauter, 1988]

Name: Lauter, Judith                    Date: 06/09/88      JL  6(3/88   TOLC( 1 4/6)
Age:   44.0 yrs    # Epochs: 48          Time: 14:22:24

Monopolar Raw Measures

|  |  | Fp1 | Fp2 | F7 | F8 | F3 | F4 | C3 | C4 | Fpz | Fz | Cz |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Absolute | Δ | 7.5 | 6.8 | 5.3 | 4.0 | 7.7 | 7.5 | 6.2 | 6.1 | 7.3 | 8.7 | 8.2 |
| Power | θ | 7.6 | 6.9 | 5.3 | 4.4 | 9.8 | 9.1 | 8.3 | 7.9 | 7.5 | 11.0 | 10.2 |
| (uU²) | α | 16.3 | 15.4 | 13.3 | 10.2 | 21.4 | 19.4 | 25.3 | 22.1 | 16.2 | 22.9 | 25.6 |
|  | β | 9.4 | 9.5 | 7.2 | 5.3 | 12.4 | 12.3 | 10.7 | 10.7 | 9.3 | 13.3 | 11.2 |
|  | T | 41.0 | 38.7 | 31.2 | 25.1 | 51.3 | 48.4 | 51.5 | 46.9 | 40.4 | 56.0 | 55.3 |
| Relative | Δ | 18.4 | 17.7 | 17.0 | 16.2 | 15.0 | 15.6 | 12.0 | 13.0 | 18.0 | 15.5 | 14.9 |
| Power | θ | 18.5 | 17.8 | 17.0 | 17.7 | 19.1 | 18.8 | 16.0 | 16.9 | 18.6 | 19.6 | 18.5 |
| (%) | α | 39.9 | 39.8 | 42.6 | 40.5 | 41.6 | 40.0 | 51.0 | 47.2 | 40.2 | 41.0 | 46.2 |
|  | β | 23.0 | 24.5 | 23.2 | 25.3 | 24.1 | 25.4 | 20.8 | 22.8 | 23.0 | 23.7 | 20.2 |
| Power | Δ | 4.7 | | 13.2 | | 1.0 | | 0.7 | | | | |
| Asymmetry | θ | 4.8 | | 8.7 | | 3.6 | | 2.2 | | | | |
| (%) | α | 2.9 | | 13.2 | | 4.9 | | 8.5 | | | | |
|  | β | -0.3 | | 5.4 | | 0.2 | | 0.1 | | | | |
| Coherence | Δ | 94.2 | | 45.7 | | 92.5 | | 85.4 | | | | |
| (%) | θ | 94.7 | | 47.1 | | 91.6 | | 78.3 | | | | |
|  | α | 96.7 | | 56.5 | | 91.4 | | 65.5 | | | | |
|  | β | 90.4 | | 46.3 | | 80.5 | | 42.7 | | | | |

|  |  | T3 | T4 | T5 | T6 | P3 | P4 | O1 | O2 | Pz | Oz |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Absolute | Δ | 3.4 | 2.7 | 2.8 | 2.3 | 4.9 | 4.6 | 2.8 | 2.1 | 5.8 | 2.1 |
| Power | θ | 3.7 | 3.7 | 3.9 | 3.0 | 6.4 | 6.2 | 3.2 | 2.8 | 7.6 | 2.9 |
| (uU²) | α | 14.2 | 11.8 | 16.6 | 10.8 | 26.3 | 21.8 | 11.9 | 10.9 | 26.2 | 10.0 |
|  | β | 6.4 | 20.7 | 8.1 | 6.1 | 9.2 | 8.5 | 5.5 | 6.5 | 9.5 | 5.8 |
|  | T | 27.8 | 38.9 | 31.6 | 22.3 | 47.0 | 41.2 | 23.5 | 22.5 | 49.1 | 20.9 |
| Relative | Δ | 12.4 | 6.9 | 9.1 | 10.3 | 10.4 | 11.1 | 12.0 | 9.6 | 11.8 | 10.2 |
| Power | θ | 13.3 | 9.5 | 12.6 | 13.5 | 13.7 | 15.1 | 13.7 | 12.7 | 15.4 | 13.9 |
| (%) | α | 51.1 | 30.3 | 52.5 | 48.4 | 56.0 | 52.8 | 50.6 | 48.6 | 53.3 | 47.8 |
|  | β | 23.0 | 53.1 | 25.7 | 27.6 | 19.6 | 20.7 | 23.6 | 28.9 | 19.3 | 27.9 |
| Power | Δ | 11.9 | | 10.9 | | 3.3 | | 12.9 | | | |
| Asymmetry | θ | 0.1 | | 13.6 | | 1.6 | | 5.6 | | | |
| (%) | α | 9.2 | | 21.1 | | 9.4 | | 4.0 | | | |
|  | β | -52.7 | | 13.6 | | 3.7 | | -0.0 | | | |
| Coherence | Δ | 30.2 | | 40.5 | | 85.8 | | 80.7 | | | |
| (%) | θ | 10.1 | | 30.2 | | 79.5 | | 84.3 | | | |
|  | α | 5.7 | | 24.2 | | 75.0 | | 79.2 | | | |
|  | β | 9.7 | | 5.1 | | 64.6 | | 74.0 | | | |

Figure 3. "Relative hemisphere advantage" plot illustrating the use of qEEG to study functional asymmetries during stimulation with complex sounds. The values plotted represent the "power asymmetry" calculation made by the Cadwell Spectrum 32 machine (cf. Fig. 2), for a single electrode pair placed over auditory cortex (T3 and T4: refer to Fig. 1 for location). Results for eleven different conditions are shown (data collection required approximately 2 hours). [Lauter, 1988]

## Complex sounds

*S = Judy Lauter*
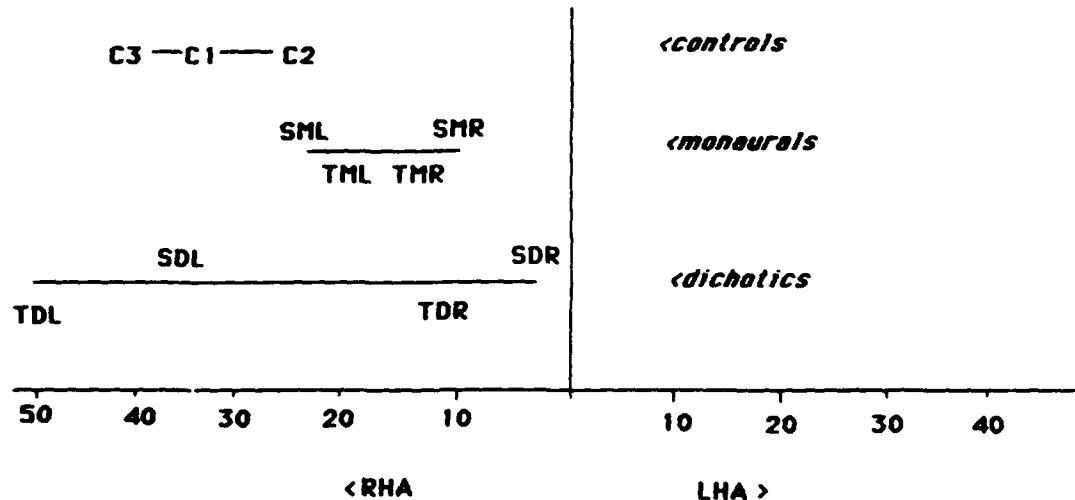
| Power asymmetry in beta EEG: |
| T3 vs. T4 |

C3 —C1—— C2                                    <controls

SML              SMR                            <monaurals
    TML  TMR

SDL                              SDR            <dichotics
TDL                        TDR

50    40    30    20    10        10    20    30    40

<RHA                              LHA >

C = Control                                    M = monaural
T = three-tone patterns                        D = dichotic
S = synthetic syllables bdgptk +a
R = right ear (attended ear in dichotic)

L = left ear (*)

order of test:
    C 1
    all syllables (m then d)
    C 2
    all tones (m then d)
    C3

Figure 4. Similar plot as in Fig. 3, showing results for a different electrode pair (F7/8--cf. Fig. 1 for location), during six _motor activation_ conditions: control, hand movement (Hand Right, Left, Both), tongue moved up and down, and "silent speech" respiration, in which the subject produced respiratory patterns appropriate to a learned prose passage without activating the larynx or upper vocal tract. [Lauter, 1988]

## Motor control

### S = Judy Lauter

Power asymmetry in beta EEG:

F7 vs. F8

C3
(9)

HL ———————— HB·HR
(-17)                (6)  (8)

Resp          Tongue
(0)            (24)

| 50 | 40 | 30 | 20 | 10 | | 10 | 20 | 30 | 40 |

< RHA          LHA >

Order of test.
   C3
   HR, HL, HB
   Tongue up down, Respiration for speech (no sound)
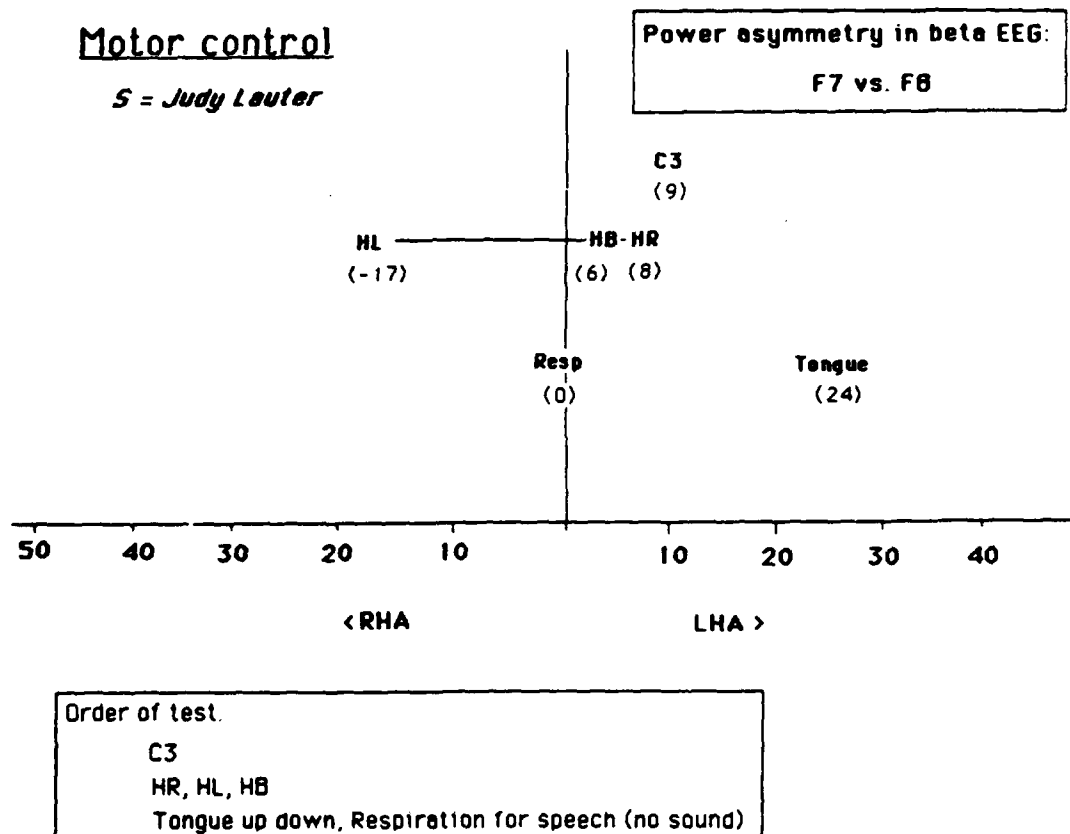
## MAGNETOENCEPHALOGRAPHY (MEG)

"Biomagnetic" fields generated by body organs were first described for the heart by Baule and McFee (1963), and later for the brain by Cohen (1968). These fields are produced as a result of electrical activity in tissue (cf. Fig. 1), and although extremely small (weaker than the earth's gravitational pull on a human body), can be detected by a device called a SQUID [(Superconducting QUantum Interference Device)]. Current design of the superconducting detectors requires that they be encased in a liquid-helium-filled cylinder known as a "dewar." This makes the monitoring interface rather bulky, and limits most SQUIDS to a single detector, though some multi-detector SQUIDS, with array sizes up to 7 detectors, have been built.

Figure 1. Schematic showing the different orientation of EEG currents and MEG fields, and the MEG detectors in the external dewar.
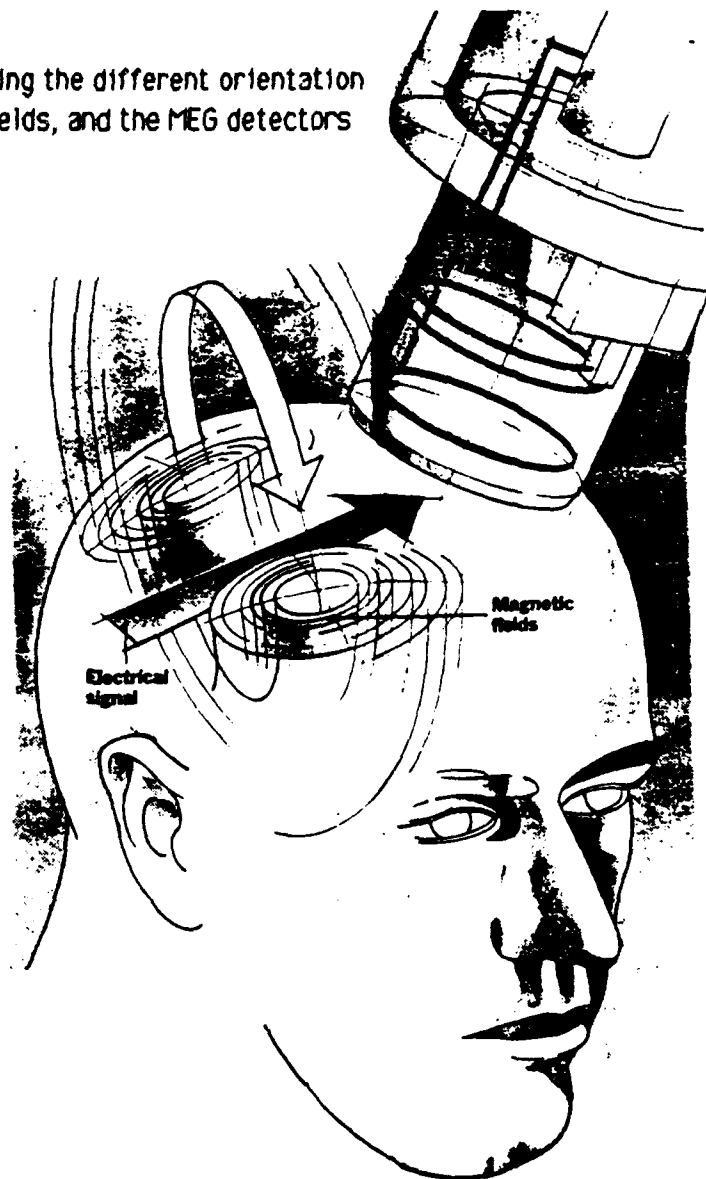
The name of MEG suggests parallels with EEG; these are most obvious in the way MEG data are collected, as averaged evoked fields (EFs), analogous to EPs.

Figure 2. Evoked-field waveforms obtained at each of 7 detectors in a 7-channel SQUID, in response to sequences of 6 syllables presented one syllable every 250 msec, with the sequence repeated once every 2 sec. [Unpublished results from a series collected by Lauter in 1987]
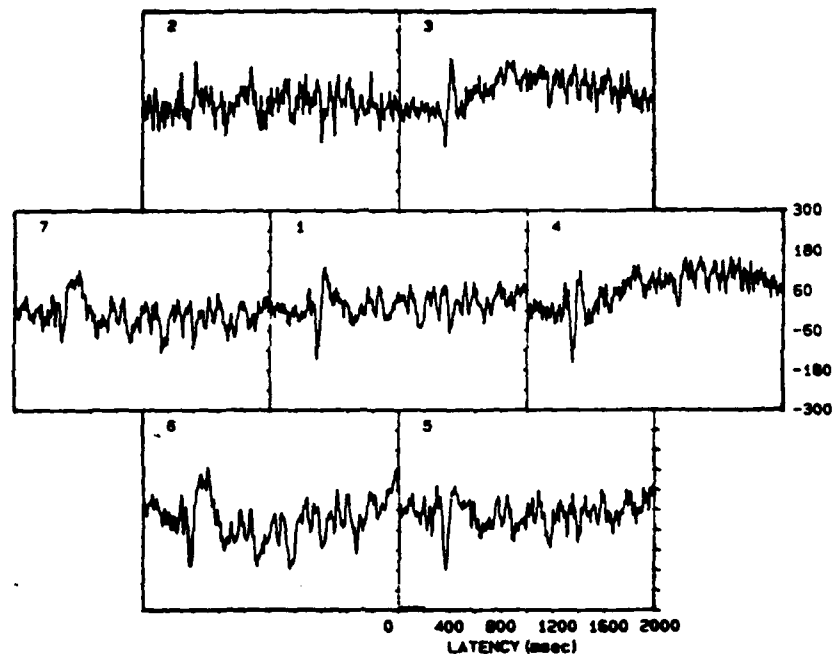
MEG does not suffer from the volume-conduction artifacts that hamper localization of EEG sources, and thus EF sources can be identified within 1 mm. This high spatial resolution has its disadvantages, however, in that the small detector arrays of current SQUIDS require multiple placements of the dewar in order to map responses over even a small area. Rapid developments in superconductor technology may within a few years radically change the shape of SQUIDS and the nature of MEG testing. Newer multi-detector arrays may be encased in helmets, allowing high-resolution topographic mapping of the entire

cortex during a single stimulation.

Subject preparation is minimal. No electrodes are used, not even to monitor eye movements, since eye movement artifacts do not affect the MEG. Some laboratories fit a flexible plastic cap on the subject's head before testing in order to record the position on the head of different placements of the dewar. Recording responses must be done with the subject and dewar in a magnetically-shielded room, where, as in MRI, no ferrous metals are tolerated. As a result, for auditory stimulation studies, transducers must be placed outside the test room, and sounds are delivered to the subject via plastic tubing. For testing, the subject lies on a table, with the dewar pressed firmly against the head over the brain region being monitored. A headrest with foam supports is used to ensure that the subject's head is kept in the same position during testing. The room is quiet and darkened. Response to one stimulus at one dewar location can be collected in 2 min. Subject preparation and data collection can be accomplished by one person, but two are preferable, one to monitor the subject and one to operate the SQUID's host computer. Analysis routines for SQUID machines are in a developmental stage, requiring the host computer, and no true "imaging" versions are available; those topographic maps that are produced are quite primitive (cf. Fig. 2). MEG testing is still at a research stage, and expenses cannot be charged to patients.

Over the past two decades, SQUID systems have been used to study brain activity in response to a variety of activation paradigms, involving both motor anc sensory stimulation (for a midpoint review, cf. Williamson & Kaufman 1981). For example, Okada et al (1982) demonstrated that left index finger flexion evoked EFs exactly over the appropriate region of motor cortex.

Figure 2. Graphs from an experiment on MEG and finger movement (Okada et al 1982). Panel A: schematic showing multiple dewar placements used to do the mapping for movement of a finger on the left hand. Panel B: EF plots demonstrating the focus of activation during left index finger flexion. The dimensions of the graph show the focus is at a point approximately 20 cm along the midline posterior to the nasion (abscissa) and 13 cm above the ear canal (ordinate) [cf. homunculus on p. 13]
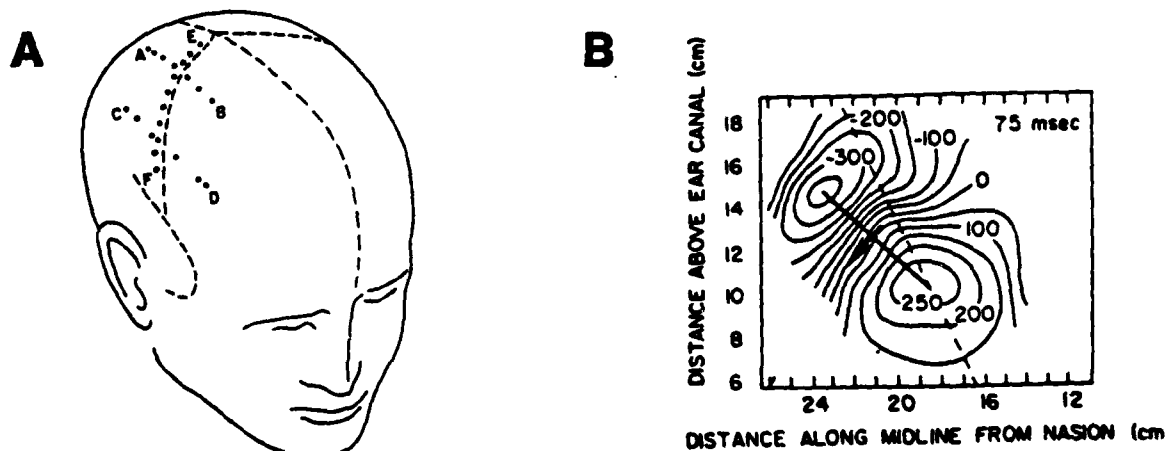
Using auditory signals, a number of characteristics of responses in human subjects have been described, including localization of the 100-ms-post-stimulus-onset portion of the (EF) response at the Sylvian fissure (e.g., Bak et al 1981), distribution of responses to different types of sounds (Reite et al 1982), nonlinear changes in EF amplitude as a function of stimulus intensity (e.g., Elberling et al 1982), and tonotopic organization (e.g., Romani et al 1982).

Figure 3. Procedures and results for experiments using MEG to study auditory activation. Panel A: schematic showing multiple dewar placements required to map auditory EFs (Elberling et al 1982). Panel B: EF plot showing a focus of responses to binaural, 32/sec clicks at approximately 1 cm anterior to the ear canal (")" on the abscissa) and 7 cm above it (from Romani et al 1982).
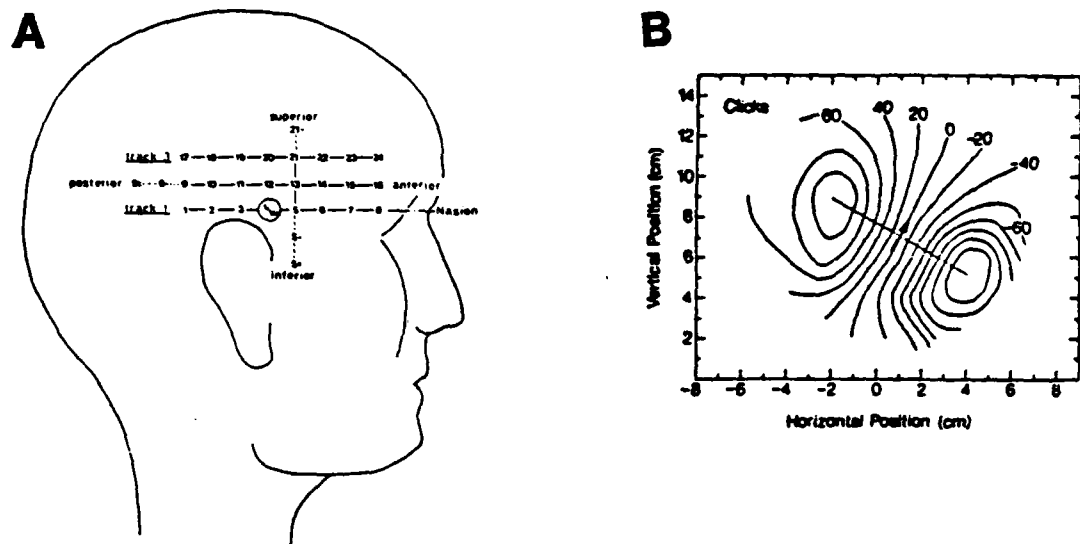


As the spatial resolution scale in the introduction suggested, MEG promises to provide the most localizable responses of any technique now available for human subjects, approaching the level of brain organization represented by cortical columns. However, since the efficiency of SQUID testing is hampered by the limitations imposed by current superconductor technology on detector-array size, the best way to use MEG in the near future may be in conjunction with another imaging method. For example, subjects might be pre-tested with PET (see next section), to determine the brain area showing maximal response to a stimulation condition, and then this region would be mapped in detail by successive SQUID dewar placements.

## POSITRON EMISSION TOMOGRAPHY (PET)

Both PET and SPECT (single photon emission computerized tomography). represent special instances of a method called autoradiography, or "taking your own X-ray." All autoradiography involves the use of isotopes introduced into a subject's cardiovascular system--so that the source of radiation is within the subject, instead of external, as in X-rays or CT--and can be used to study any part of the body. Depending on the isotope and the compound it is used to "label" (e.g., glucose labelled with an isotope of carbon), autoradiography can used to study metabolism (e.g., uptake of dopamine in areas of the brain related to Parkinson's Disease), blood supply (e.g., volume and movement of blood within the heart), and/or function (e.g., using blood flow as a measure of differential activity in separate areas of the brain).

Methods such as PET and SPECT grew out of experiments in invasive autoradiography, where isotopes with radiation products too weak to penetrate bone are used to study nonhuman animals. In these experiments, when the goal is to study the brain, the animal must be killed at the end of the experiment, and the brain removed from the skull in order to study distribution of the isotope. The development of noninvasive autoradiography for use in humans, depended on the introduction of isotopes which gave rise to photons, which are powerful enough to penetrate the skull. Both PET and SPECT are ways to externally monitor photons released from isotopes within the body; isotopes that can be used include radioactive forms of flourine, iodine, carbon, and oxygen.

Positron-emitting isotopes such as oxygen-15 give off positrons, which collide with electrons in their immedite vicinity in "annihilation events," yielding pairs of gamma rays radiated at 180 degrees to each other. Thus a circle of detectors, such as that comprising the PETT VI device (Ter-Pogossian et al 1982) can detect these "coincidences" of radiated photon pairs, and the detections are signalled to a computer which keeps track of the distribution of the events so monitored (cf. Figs. 1 and 2).

Figure 1. The "annihilation event" that is at the heart of the PET-scan, which makes it a noninvasive autoradiography tool ideal for use with human subjects.
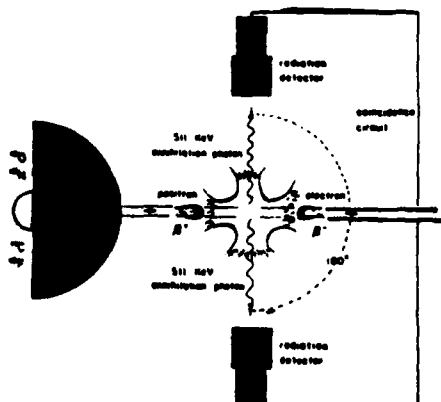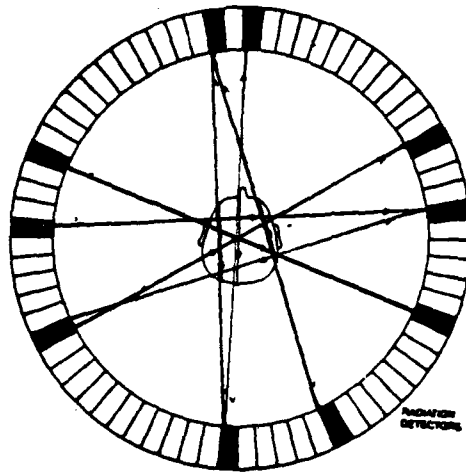
Figure 2. Design of the PETT-VI, a cylinder of detectors that can provide a three-dimensional view of the entire cranial CNS in one 40-sec scan.

The distribution of decay events can be used to generate color-coded tomographic maps, showing regions of greater or lesser isotope concentration, interpreted as regions of greater or lesser blood flow, implying (because neuronal function requires an ongoing supply of glucose and oxygen) regions of greater or lesser brain activity (cf. Raichle et al 1976).

The usefulness of the different methods for studying brain activity is determined by 1) the type of machine used, 2) the means of administering the isotope, and 3) the half-life of the isotope. Some machines can only monitor a small area of brain during a single scan, while others can produce for each scan a complete three-dimensional series of images representing the whole brain. How the isotope is administered is important with regard to the invasiveness of the procedure. In older PET studies, isotopes were often injected into a carotid artery. Administration method also affects the reliability of the measurement: e.g., the initial amount of radioactivity in an isotope injected as a bolus into an arm vein can be measured more reliably than that in an isotope which is inhaled.

Perhaps most important is the half-life of the isotope. Half-life determines the dosage of radioactivity given to the subject while the isotope is in the body, and this directly influences the type of experiments that can be done as well as the interpretability of results. Isotopes with long half-lives (e.g., several hours) deliver a larger dosage of radioactivity to the subject. As a result, each subject can be tested only once, thus making it impossible to do experiments where responses to several conditions can be compared in the same subject. In addition, these isotopesand require quite long

scan times, e.g., 45 minutes, and it is difficult to believe that the brain is "doing the same thing" during the entire time of the scan. Many of the PET experiments that have been published regarding speech and hearing applications are based on use of a type of glucose called fluoro-deoxyglucose (FDG), labelled with an isotope of carbon with a long half-life. The reported results regarding loci of greatest activation during speaking, or asymmetries in reponses to certain types of sounds must be considered in the context of these problems associated with PET testing based on long half-life isotopes, and interpreted with great caution.

The most useful approach to date for studying human brain function is represented by a cylindrical, multi-detector device (such as the PETT VI, or the SuperPet), combined with use of a short-half-life isotope such as oxygen-15, with a half-life of approximately 2 min. This combination makes it possible to take whole-brain scans in as little as 40 seconds, and the dosage to the subject is so low that 8 scans can be run in a single test session, and three such test sessions can be given to the same subject in the same year.

Programs for data reduction provide a number of approaches to studying functional responses as they appear on PET images. For example, one can: 1) obtain quantitative estimates of rCBF within a selected region, at any point in the horizontal or vertical plane; 2) compare a slice from a scan taken under control conditions with the same slice from a scan taken under stimulation conditions, and generate a new map showing regions of greater or lesser change from control to stimulation; 3) compare the two hemispheres of a single slice, and generate a third type of map showing regions in one hemisphere that are more or less different from mirror-image regions on the opposite side. Images can also be reconstructed in the sagittal or coronal planes, and the same types of analysis applied. Routines are also available for combining measurements taken from lateral skull radiographs of each subject, and the PET images, in order to estimate the anatomical localization of regions of interest (Fox et al 1985).

Procedures include insertion of an intravenous cathether into an arm vein, fitting the subject with a face mask to hold the head in place during testing, and, for each scan, injection of a bolus of oxygen-15-labelled water. Subject preparation requires 30-60 min., and each scan requires 40 sec. for data collection, with intervening periods of approximately 15 min. to allow residual activity to diminish before injection of another bolus for testing the next condition. Testing is done in a quiet, darkened room, with the subject reclining with head inside the cylinder of detectors. PET testing is personnel-intensive, requiring at least five individuals' participation during scanning: two cyclotron operators working in a nearby area, a person in the test room to monitor the subject and administer test conditions, a second in the test room to inject the isotope, and one to operate the PET machine. Each PET system has its own unique arrangements for data analysis; some require use of the data-collection system, while others make use of minicomputer-based systems for satellite analysis. One laboratory has developed

analysis routines for satellite Mac II microcomputers. At this date, all PET testing is done on a research basis; there are no third-party reimbursements for this type of brain analysis.

The first attempt to measure changes in regional cerebral blood flow or metabolism (rCBF/M) in human brains in response to a behavioral stimulus was reported by Sokoloff et al (1955). In these experiments, subjects were asked to solve a set of arithmetic problems while being monitored for arterio-venous differences in the level of inhaled nitrous oxide gas. Since this time, a number of methods for monitoring rCBF/M have been developed, involving more or less noninvasive procedures, and exploited for studying both normal and abnormal human brains, including brain responses when tested under a variety of activation conditions, ranging from single-finger movement to visual pattern recognition (for a review, cf. Raichle, 1987). We have reported controlled studies using auditory stimulation that demonstrate: 1) changes in auditory-cortical response to parameters such as rate, level, and ear-of-presentation (Lauter et al 1983: Fig. 3); 2) tonotopic organization in human primary auditory cortex (Lauter et al 1985a: Fig. 4); and 3) simultaneous activation of cortical and subcortical auditory centers during a single stimulation (Lauter et al 1985b).

Although a number of papers have been published describing brain responses monitored with PET during activation conditions related to speech and hearing (e.g., while subjects are speaking, or listening to stories or hearing different types of music), the results of almost all these experiments are qualified by the methods used to do the scans, such as long scan times, lack of within-subject comparisons, and poor control over activation details such as phonemes produced or the levels of test sounds. As with any device, the potential of PET for exploring communication disorders will not be realized unless studies involving rigorous control of all test variables are done; lacking such control, the 'pretty pictures' produced by PET remain uninterpretable.

Figure 3. Quantitative results of controlled parametric studies using PETT VI and oxygen-15. All graphsare based on an index of brain activity consisting of percent change in blood flow from control to test condition. Panels A, B, and C show blood-flow changes in selected motor, visual, and auditory areas, respectively, as a function of rate of activation. Panels D and E illustrate two other parameters studied with auditory stimulation, level and ear-of-presentation. Graph B is from a study with 6 subjects; the other panels represent single-subject data. [From Lauter et al 1983; visual data from Fox and Raichle 1984]
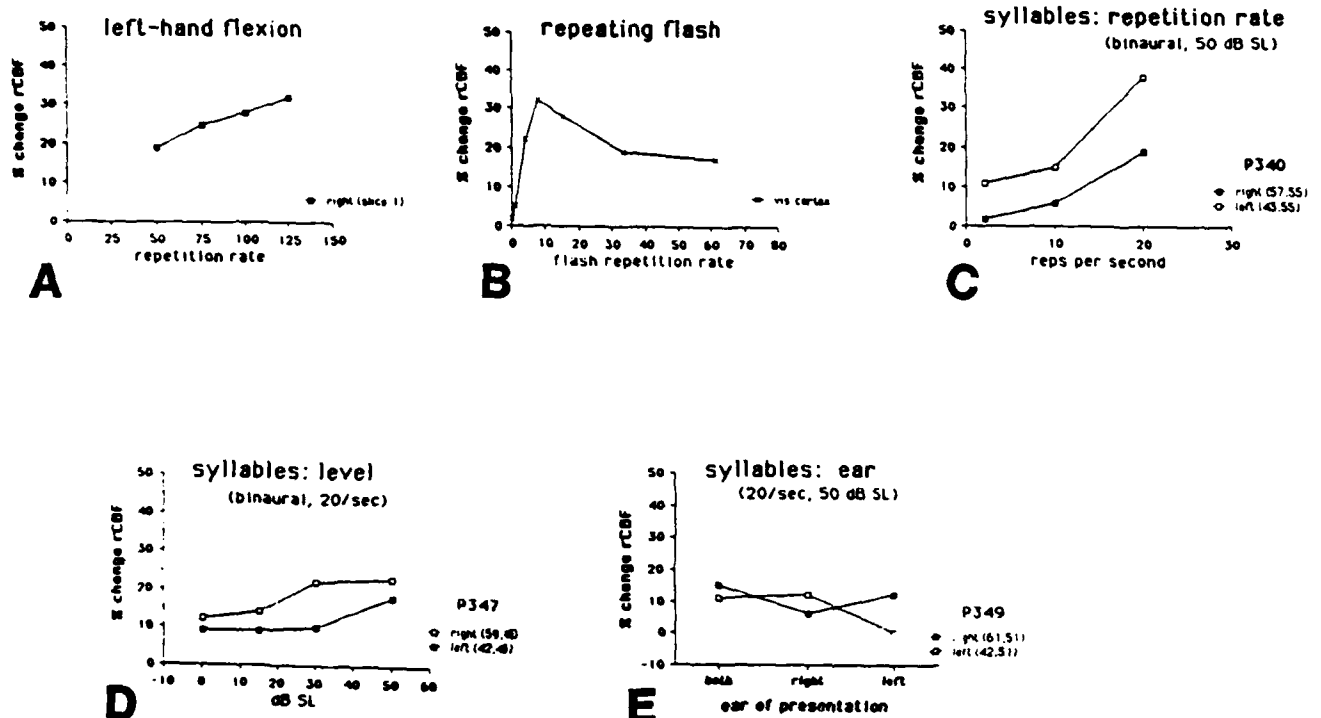
Figure 4. Data demonstrating tonotopic organization in human primary auditory cortex, collected using PETT VI and oxygen-15. Panel A is a schematic showing the relative location in PET images of regions responding best to each of two frequencies of pulsed pure tones presented monaurally to the right ear: these regions were located in left-hemisphere primary auditory cortex for each of the five subjects. Regions responding best to the 500-Hz tone are indicated by the filled symbols, regions responding best to the 4 kHz tone are indicated by the open symbols. Panel B provides a comparison of the orientation of responses (lower frequencies represented more anteriorly and laterally, higher frequencies more posteriorly and medially) to similar tones, collected from cells in primary auditory cortex of rhesus monkeys [PET results based on data from Lauter et al 1985].



rCBF response areas in 5 humans  A

● response to 500 Hz tone
□ response to 4 kHz tone
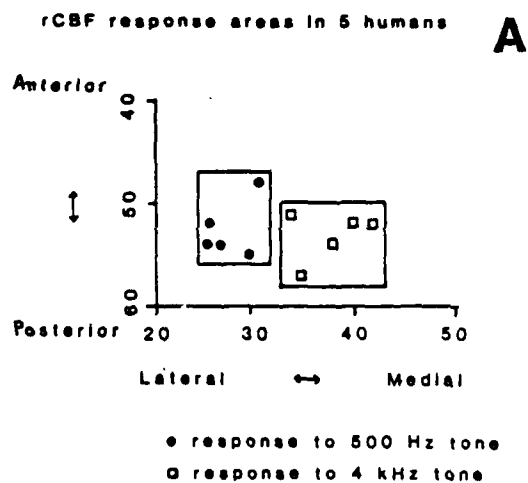
Superior temporal plane in monkey  B
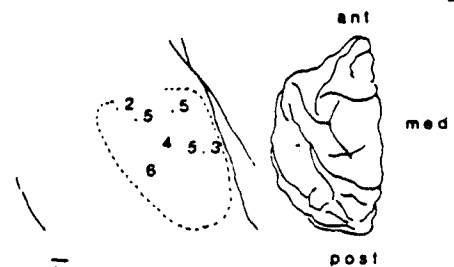
--from Merzenich and Brugge (1973)

CONCLUSION

The resolution scales reproduced in the Introduction, comparing the different noninvasive devices currently available, illustrate that these devices already provide a range of access to activity in the living human brain, with observation times stretching from 40 sec to hours, and from a behavioral level of spatial detail to one approaching macrounits of neuronal organization. It is not at all clear at this time which of the machines will be most useful for studying issues related to human communication, either in healthy subjects or in individuals with different types of dysfunction. There are trade-offs in spatial and temporal acuity of each machine, as well as possibilities for using methods in conjunction to tailor testing to the characteristics of individual subjects.

The Coordinated Noninvasive Studies (CNS) Project at University of Arizona has been designed to take the first steps toward establishing how these noninvasive devices might most efficiently be used to study human brain function. Initial stages of the Project will focus on brain asymmetries, including those related to speech production and auditory perception. Asymmetries are easily studied with all the devices discussed above, and represent expressions of basic principles of brain organization such as specialization of function with regard to "side of space," as well as organization for controlling both hierarchical and parallel processing. For this project, several subjects will be tested on all the devices named above. First, each subject will be examined using psychophysical methods to determine behavioral profiles of "relative asymmetries" for sets of stimuli (such as ear advantages for speech vs. nonspeech sounds: cf. Lauter 1982, 1983, 1984). Next each subject will receive an MRI scan so that anatomical asymmetries can be measured, and will then undergo successive tests on each of the physiological devices while being presented with the same stimuli used in the behavioral tests. The asymmetry patterns yielded by each method will then be examined to determine their correlations with each other, and with the behavioral data.

With regard to clinical applications, results can most immediately provide guidelines for studying a variety of neurogenic communication disorders, ranging from "central auditory disorders" associated with subcortical pathology, to stuttering and aphasia. Future adaptations in test selection and methodology will allow us to study individuals with more ambiguously defined dysfunctions, such as autism and dementia, and to obtain detailed information regarding brain function even in subjects who cannot actively participate in testing. It is to be expected that the new insights provided by these noninvasive devices and their successors will revolutionize our concepts of the human brain, and how it works to achieve communication.

REFERENCES


Bak C, B Kofoed, J Lebech, K Saermark, C Elberling (1981) Auditory evoked magnetic fields from the human brain. Source localisation in a single-dipole approximation. Phys Lett 82A: 57-60.


Baule GM & R McFee (1963) Detection of the magnetic field of the heart. Amer Heart J 66: 95-96.


Berlin CI, LJ Hood, P Allen (1984) Asymmetries in evoked potentials. In CI Berlin (Ed), Hearing science. College-Hill: San Diego; pp. 461-478.


Brant-Zawadzki M & D Norman (Eds) (1987) Magnetic resonance imaging of the central nervous system. Raven Press: NY.


Brazier MAB (1977) Electrical activity of the nervous system. (4th ed) Pitman Medical: London.


Chui HC & AR Damasio (1980) Human cerebral asymmetries evaluated by computed tomography. J Neurol Neuropsych Psychiat 43: 873-878.


Cohen D (1968) Magnetoencephalography: Evidence of magnetic fields produced by alpha-rhythm currents. Science 161: 784-786.


Cohen SM (Ed) (1987) Physiological NMR spectroscopy: from isolated cells to man. Ann NY Acad Sci 508: 1-537.


Courchesne E (In press) Cerebellar changes in autism. In J Swann and A Messer (Eds) Disorders of the developing nervous system: Changing views on their origins, diagnoses and treatments. Alan R. Liss, Inc: NY.


Courchesne E, JR Hesselink, TL Jernigan, R Yeung-Courchesne (1987) Abnormal neuroanatomy in a nonretarded person with autism. Arch Neurol 44: 335-341.


Davis H (1976) Principles of electric response audiometry. Ann Otol Rhinol Laryngol 85: Supp 28.

Duffy FH, JL Burchifiel, CT Lombroso (1979) Brain electrical activity mapping
(BEAM): a method for extending the clinical utility of EEG and evoked potential data. Ann
Neurol 5: 309-321.

Elberling C, C Bak, B Kofoed, J Lebech, K Saermark (1982) Auditory magnetic fields
from the human cerebral cortex: Location and strength of an equivalent current dipole.
Acta Neurol Scand 65: 553-569.

Fox PT & ME Raichle (1984) Stimulus rate dependence of regional cerebral blood flow
in human striate cortex, demonstrated by positron emission tomography. J Neurophysiol
51: 1109-1120.

Fox PT, JS Perlmutter, ME Raichle (1985) A stereotactic method of anatomical
localization for positron emission tomography. J Comp Assist Tomog 9: 141-153

Galaburda AM, M LeMay, TL Kemper, N Geschwind (1978) Right-left asymmetries in the
brain. Science 199: 852-856.

Gevins AS & RE Schaffer (1980) A critical review of electroencephalographic (EEG)
correlates of higher cortical functions. CRC Crit Rev Bioengineering 4: 113-164.

Geschwind N & W Levitsky (1968) Human brain: Left-right asymmetries in temporal
speech regions. Science 161: 186-187.

John ER (1977) Neurometrics: clinical applications of quantitative electrophysiology.
In Functional neuroscience, vol. 2. Erlbaum: Hillsdale NJ.

Kertesz A, SE Black, M Polk, J Howell (1986) Cerebral asymmetries on magnetic
resonance imaging. Cortex 22: 117-127.

Lauter JL (1982) Dichotic identification of complex sounds: Absolute and relative ear
avantages. JASA 71: 701-707.

Lauter JL (1983) Stimulus characteristics and relative ear advantages: A new look at
old data. JASA 74: 1-17.

Lauter JL (1984) Contralateral interference and ear advantages for identification of three-element patterns. Brain & Cognition 3: 259-280.

Lauter JL (1988) Functional-activation asymmetries in normal humans studied with quantitative EEG (qEEG): first tests in the CNS Project. Presented to Acoustical Society of America, Honolulu.

Lauter JL & RL Loomis (1986) Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability. I. Absolute latencies of brainstem vertex-positive peaks. Scand Audiol 15: 167-172.

Lauter JL & RL Loomis (in press) Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability. II. Amplitudes of brainstem vertex-positive peaks. Scand Audiol.

Lauter JL, P Herscovitch, ME Raichle (1983) Human auditory cortex: a preliminary report on studies using positron emission tomography (PET), presented to Acoustical Society of America, Cincinnati. Abstract: J Acoust Soc Amer 73: S60.

Lauter JL, P Herscovitch, C Formby, ME Raichle (1985a) Tonotopic organization in human auditory cortex revealed by positron emission tomography. Hearing Res 20: 199-205.

Lauter JL, Herscovitch, ME Raichle (1985b) Human auditory central nervous system: visualization of cortical and subcortical centers with regional cerebral blood flow measurements on the PETT-VI, to Acoustical Society of America, Austin. Abstract: J Acoust Soc Amer 77: S94.

Levine RB & PM McGaffigan (1983) Right-left asymmetries in the human brain stem: auditory evoked potentials. Electroenceph Clin Neurophysiol 55: 532.

Okada YC, SJ Williamson, L Kaufman (1982) Magnetic field of the human sensorimotor cortex. Internat J Neurosci 17: 33-38.

Pieniadz JM & MA Naeser (1984) Computed tomographic scan, cerebral asymmetries and morphologic brain asymmetries. Arch Neurol 41: 403-409.

Raichle ME (1987) Circulatory and metabolic correlates of brain function in normal humans. In F Plum (Ed) Handbook of physiology. The nervous system V. Higher functions of the brain, Part 2. American Physiological Society: Bethesda MD; pp 643-674.

Raichle ME, RL Grubb, MH Gado, JO Eichling, MM Ter-Pogossian (1976) Correlation between regional cerebral blood flow and oxidative metabolism. Arch Neurol 33: 523-526.

Ratcliff G, C Dila, L Taylor, B Milner (1980) The morphological asymmetry of the hemispheres and cerebral dominance for speech: A possible relationship. Brain Lang 11: 87-98.

Reite M, JT Zimmerman, JE Zimmerman (1982) MEG and EEG auditory responses to tone, click and white noise stimuli. Electroenceph Clin Neurophysiol 53: 643-651

Romani GL, SJ Williamson, L Kaufman, D Brenner (1982) Characterization of the human auditory cortex by the neuromagnetic method. Exp Brain Res 47: 381-393.

Rubens AB, MW Mahowald, JT Hutton (1976) Asymmetry of the lateral (Sylvian) fissures in man. Neurol 26: 620-624

Sokoloff L, R Mangold, RL Wechsler, C Kennedy, SS Kety (1955) The effect of mental arithmetic on cerebral circulation and metabolism. J Clin Invest 34: 1101-1108.

Ter-Pogossian MM, DC Ficke, JT Hood, Sr., M Yamamoto, NA Mullani (1982) PETT VI: Positron emission tomography utilizing cesium fluoride scintillation detectors. J Comp Assist Tomogr 6: 125-133.

Teszner D, A Tzavaras, J Gruner, H Hecaen (1972) L'asymetrie droit-gauche du planum temporale: A propos de l'etude anatomique de 100 cerveaux. Rev Neurologique 126: 444-449.

Valk J, C MacLean, PR Algra (1985) Basic principles of nuclear magnetic resonance imaging. Elsevier: Amsterdam.

Wada J, R Clarke, A Hamm (1975) Cerebral hemispheric asymmetry in humans. Arch Neurol 32: 239-246.

Williamson SJ & L Kaufman (1981) Biomagnetism. J Magnetism Magnetic Materials 22: 129-201.

Witelson SF (1982) Bumps on the brain: right-left anatomic asymmetry as a key to functional lateralization. In S Segalowitz (Ed), Language functions and brain organization. Academic: NY; pp. 117-144.

Witelson SF & W Pallie (1973) Left hemisphere specialization for language in the newborn: Neuroanatomical evidence of asymmetry. Brain 96: 641-647.

Functional-activation asymmetries in normal humans studied with quantitative EEG
(qEEG): first tests in the CNS Project. Judith L. Lauter (University of Arizona, Tucson
AZ 85721)·Presented to ASA, Honolulu, November 1988. Abstract: J Acoust Soc Amer
84: S57.

## ABSTRACT

Our work with behaviorally-defined asymmetries such as relative ear advantages has
led to the CNS Project, designed to apply noninvasive tech- niques such as psychophysics,
EPs, qEEG, MEG, MRI, and PET to study human brain responses during functional
activation. Preliminary results of a qEEG test series involving both auditory and hand-
movement conditions indicate that qEEG power asymmetry patterns reflect at least two
types of asymme- try organization: 1) "side of space," e.g., right-hand movement
elicits a power asymmetry favoring the left hemisphere, and v.v; and 2) asymmetries
based on "higher-level" principles of organization, e.g., coordination during bilateral
hand movement, or differential activation based on the physical characteristics of test
sounds. As with behavioral patterns of relative ear advantages, qEEG shows individual
differences in detail but group agree- ments in overall patterns of response. To our
knowledge, this is the first report of qEEG used in this way for studying functional
activation in healthy human subjects, and illustrates its potential usefulness for studying
human neurophysiology. [Supported in part by AFOSR]

## INTRODUCTION

For more than a century, electroencephalography (EEG) has been employed for
studying the human brain, by means of both real-time and averaged forms of scalp-
recorded potentials. Spectral analysis of ongoing EEG (quantitative EEG, or qEEG) has
provided detailed information supporting research on "cognitive processing," including
questions related to cerebral asymmetries (for reviews, cf. Gevins & Schaffer 1980,
Gevins 1984, Glass 1984).

However, surprisingly little is known about the contribution made by "other-than-
cognitive" (cf. Gevins 1980) processes to the patterns of EEG activity. Careful study of
the effects of relatively simple variables such as rate and level of stimulation; or basic
factors related to cerebral asymmetries, such as contralateral vs. ipsilateral
representation and influence of stimulus characteristics, may provide primary
information regarding brain organization and function, and may even help account for
results observed in experiments focused on more "mental" operations.

## TESTING

We referred to our research on dichotic listening (Lauter 1982, 1983, 1984) to
design an experimental paradigm for use with qEEG. Subjects are pre-tested using

behavioral methods, to familiarize them with monaural and dichotic identification of two sound sets: 1) six synthetic stop consonant-vowel syllables (coded as "S" in the figures below), and 2) six three-note patterns made with three pure tones set at 1440, 1480, and 1520 Hz, with the 25-msec tones set at 200 ms between onsets (coded as "T"). Broad-band spectrograms of the syllables and schematics of the tone patterns are presented in Fig 1. Identification responses on a total of six 36-trial blocks per sound per subject are collected with ear-of-report alternating from block to block; experimental blocks are approximately 5 min in length.

Fig. 1. Two sets of test sounds: synthetic nonsense stop- consonant-vowel syllables, and three-tone patterns.
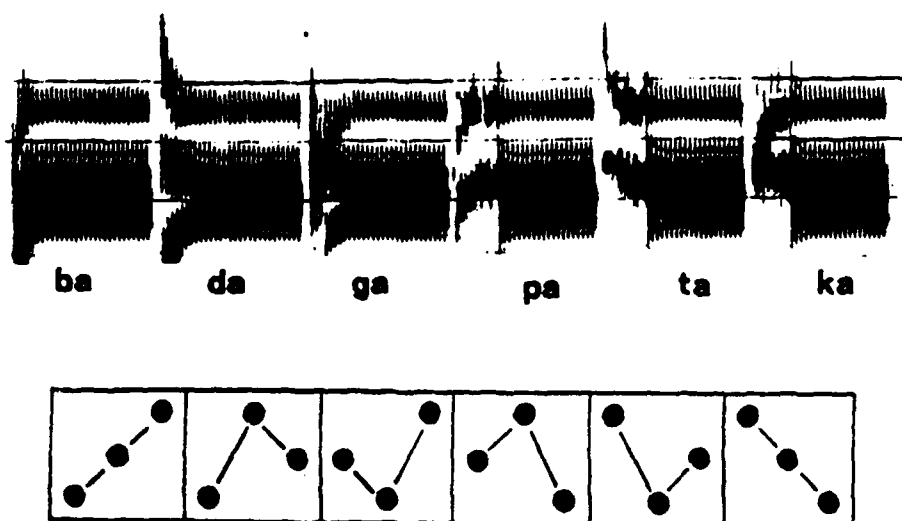


Testing for both sound sets requires a total of four hourly sessions per trained subject; different sounds are tested on different days. For two of our subjects, who were trained listeners, complete test series were collected for both sound sets (Subjects JL and CB). The other two subjects (DW and JM) found dichotic identification of the tone patterns quite difficult, and were unable to remain in the experiment long enough to achieve better-than-chance performance with dichotic presentation. The ear advantages obtained for all listeners are displayed in Fig 2.

Fig. 2 Ear advantages shown by eacn of the 4 subjects on the test sounds; JM and DW were unable to identify the tone patterns when presented dichotically.
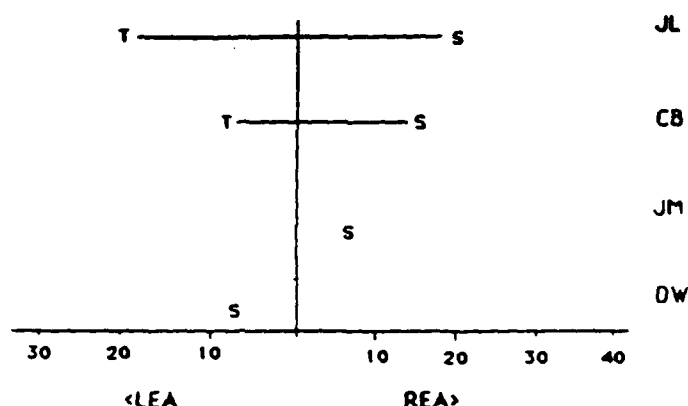


After behavioral testing was completed, individuals were scheduled for qEEG testing. Preparation (Fig. 3: color picture of a subject and a qEEG machine, not included for this MS) includes fitting of an electrode cap, with leads connected to a Cadwell Spectrum 32 qEEG system, with capabilities for multi-channel data collection and spectral analysis. Electrodes are placed at 8 locations over each hemisphere, and 5 locations along midline, according to the 10-20 system; potentials at all locations are referenced to linked earlobes. When impedance for each of the leads is less than 8 ohms, testing is begun.

The schedule of conditions is shown in Table I. A time base similar to that used in the behavioral testing is used, with 5 min of EEG collected during each test condition. Note that each test session concludes with a brief set of blocks involving motor activation. Throughout, the subject reclines in a comfortable chair in a quiet, darkened room. Test sounds are played via a stereo cassette recorder through stereo earphones. During monaural stimulus conditions, subjects are told to attend to the ear of input; during dichotic conditions, they are told to attend to the ear targeted for that condition in the same way done for the behavioral tests previously. We do not ask for score-able identification performance during the EEG testing, in order to avoid movement artifacts. Trained subjects report that it is easy to perform this "mock" dichotic listening. The qEEG results suggest that the two trained listeners here were in fact successfully replicating processing patterns used in the behavioral testing.

Table I. Conditions tested per session
1. Control (no activation)
2. Synthetic syllables in left ear
3. Synthetic syllables in right ear
4. Synthetic syllables dichotically, attend to right ear
5. Syllables dichotically, attend to left ear
6. Control
7. Tone patterns in right ear
8. Tone patterns in left ear
9. Tone patterns dichotically, attend to left ear
10. Tone patterns dichotically, attend to right ear
11. Control
12. Preferred hand flexion, 1/sec
13. Opposite hand flexion, 1/sec
14. Bilateral hand flexion, 1/sec
15. Control

DATA ANALYSIS

Data were analyzed off-line. From each 5-min EEG record, 36 2.5-sec artifact-free epochs were selected by eye (cf. Fig. 4: a color figure, showing the EEG waveform display; not included in this MS). The Cadwell Spectrum 32 then averaged the selected epochs, performs spectral analysis according to 4 EEG bandwidths (see Table II), and displayed the results in terms of the parameters shown in Table II. From each table representing each subject tested on each condition, a single number is chosen: 1) for the auditory conditions, the value used is the beta power asymmetry comparing temporal-lobe electrode locations T3 and T4; 2) for the hand-movement conditions, beta power asymmetry comparing F7 vs. F8 is used.

Table II. Sample data table calculated for each test condition, showing values for 4 qEEG parameters by electrode location and bandwidth. (Bandwidth ranges shown at bottom).

Name: Lauter, Judith
Age: 44.0 yrs    # Epochs: 40

Date: 06/07/90
Time: 10:10:10

Monopolar Raw Measures

| | | Fp1 | Fp2 | F7 | F8 | F3 | F4 | C3 | C4 | Fpz | Fz | Cz |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Absolute | Δ | 6.2 | 6.8 | 4.4 | 3.9 | 7.0 | 6.5 | 6.1 | 6.8 | 6.1 | 7.6 | 7.6 |
| Power | θ | 5.9 | 6.6 | 4.3 | 3.5 | 7.7 | 7.1 | 7.5 | 6.8 | 5.9 | 8.4 | 6.2 |
| (uV²) | α | 8.6 | 8.2 | 6.6 | 5.3 | 10.3 | 10.1 | 15.2 | 12.0 | 8.6 | 11.5 | 11.4 |
| | β | 8.6 | 8.0 | 6.2 | 4.6 | 11.6 | 10.2 | 9.2 | 8.7 | 8.4 | 11.7 | 9.4 |
| | T | 29.4 | 27.7 | 21.6 | 17.5 | 36.7 | 34.0 | 38.1 | 33.6 | 29.1 | 39.4 | 36.8 |
| Relative | Δ | 21.1 | 21.0 | 20.3 | 22.6 | 19.2 | 19.2 | 16.0 | 17.5 | 21.1 | 19.4 | 20.0 |
| Power | θ | 20.3 | 20.4 | 20.1 | 20.2 | 20.9 | 20.8 | 19.7 | 20.4 | 20.3 | 21.3 | 22.3 |
| (%) | α | 29.2 | 29.7 | 30.7 | 30.8 | 28.1 | 29.6 | 39.9 | 35.9 | 29.5 | 29.3 | 31.2 |
| | β | 29.2 | 28.8 | 28.8 | 26.3 | 31.6 | 30.2 | 24.3 | 26.1 | 28.9 | 29.8 | 25.5 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Power | Δ | 3.1 | | 4.9 | | 3.5 | | 1.8 |
| Asymmetry | θ | 2.6 | | 10.1 | | 4.0 | | 4.4 |
| (%) | α | 2.0 | | 10.5 | | 1.1 | | 11.4 |
| [L-R] | β | 3.6 | | 14.8 | | 6.0 | | 2.6 |
| Coherence | Δ | 95.1 | | 48.1 | | 93.1 | | 88.3 |
| (%) | θ | 92.8 | | 37.3 | | 89.0 | | 77.3 |
| | α | 95.0 | | 42.2 | | 86.5 | | 42.6 |
| | β | 91.5 | | 48.4 | | 79.5 | | 32.7 |

| | | T3 | T4 | T5 | T6 | P3 | P4 | O1 | O2 | Pz | Oz |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Absolute | Δ | 2.9 | 2.8 | 3.1 | 2.6 | 5.2 | 4.9 | 3.8 | 2.7 | 6.1 | 2.7 |
| Power | θ | 3.5 | 3.0 | 3.8 | 2.7 | 6.8 | 5.6 | 3.2 | 2.8 | 6.6 | 2.9 |
| (uV²) | α | 8.4 | 7.0 | 10.0 | 5.9 | 13.3 | 10.6 | 6.5 | 6.5 | 11.8 | 5.8 |
| | β | 4.7 | 9.7 | 7.4 | 4.6 | 8.4 | 7.1 | 7.4 | 6.1 | 8.2 | 5.8 |
| | T | 19.6 | 22.7 | 24.4 | 15.8 | 33.0 | 28.2 | 21.1 | 18.3 | 32.8 | 17.3 |
| Relative | Δ | 14.8 | 12.6 | 12.7 | 16.4 | 15.8 | 17.5 | 18.3 | 14.8 | 18.7 | 15.6 |
| Power | θ | 18.1 | 13.4 | 15.6 | 17.3 | 18.1 | 19.8 | 15.3 | 15.7 | 20.1 | 16.7 |
| (%) | α | 42.9 | 31.0 | 41.0 | 37.1 | 40.4 | 37.5 | 31.0 | 35.9 | 35.9 | 33.6 |
| | β | 24.1 | 42.8 | 30.5 | 28.9 | 25.5 | 25.0 | 35.1 | 33.4 | 25.0 | 34.0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Power | Δ | 8.4 | | 8.8 | | 2.7 | | 12.7 |
| Asymmetry | θ | 7.5 | | 16.1 | | 3.3 | | 5.8 |
| (%) | α | 8.9 | | 25.9 | | 11.3 | | -0.1 |
| [L-R] | β | -34.4 | | 23.6 | | 8.6 | | 9.6 |
| Coherence | Δ | 35.3 | | 63.3 | | 90.5 | | 83.3 |
| (%) | θ | 12.9 | | 44.9 | | 83.4 | | 87.6 |
| | α | 16.4 | | 15.3 | | 66.3 | | 75.6 |
| | β | 12.2 | | 5.9 | | 60.4 | | 67.3 |

Δ = 1.5 - 3.5 Hz
θ = 3.5 - 7.5 Hz
α = 7.5 - 12.5 Hz
β = 12.5 - 20 Hz

RESULTS

Results for the auditory conditions are presented In Fig. 5, using a format adapted
from the EA graph of Fig. 2. The four panels of Fig. 5 present T3/T4 beta-asymmetry
values for 4 individual subjects tested on 3 control and 8 auditory conditions. Data for
control, monaural, and dichotic conditions are plotted on separate rows. Behavioral ear
advantages for each subject for the 2 sound sets are indicated at the top of each qEEG plot.
The qEEG results indicate evidence of 2 types of asymmetry: 1) one based on "side of
space" in that attention to right vs. left ear results in opposite asymmetries; and 2) an
asymmetry based on type of sound, in that attention to syllables vs. tones results in
opposite asymmetries. There are also interactions between the two types of asymmetry,
such that right-ear syllables tend to evoke one extreme of asymmetry and left-ear tones
the opposite extreme.

Fig. 5 qEEG asymmetries: Complex sounds
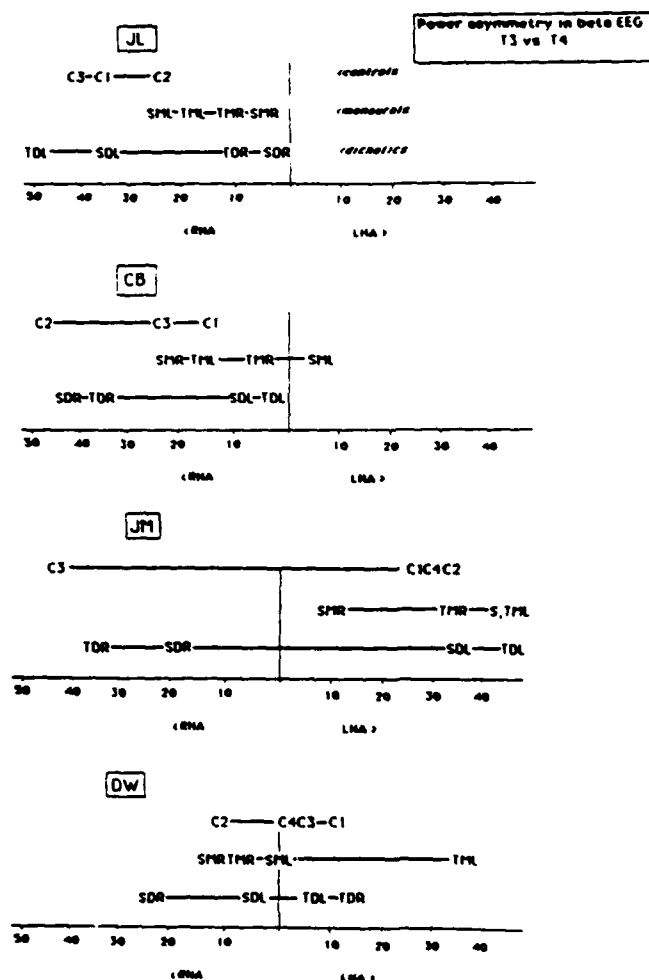
Based on the results for the subjects shown in Fig. 5, it cannot be said that the observed asymmetries always reflect predominantly contralateral activation, although this would be the predicted result. Only Subject JL shows a trend toward the expected "contralateral effect," with greater right-hemisphere asymmetries (RHA) for left-ear input/attention, and vice versa. Notice, however, that even for JL, the patterns of asymmetry are articulated in terms of modulations of RHA-- *none of the auditory conditions shows an actual left-hemisphere asymmetry for this subject.* This observation is in contrast to the behavioral results shown in Fig. 2, where the syllables evoked a 20% right-ear advantage, and the tones evoked a 20% left-ear advantage. The qEEG asymmetries suggest that JL's behavioral results may indeed reflect changes in *relative* hemisphere activation, but these are changes which occur in the context of a continuing processing predominance favoring the right hemisphere.

All of the other subjects show what must be interpreted as "ipsilateral" patterns of activation, with right-ear input/attention and syllables evoking a greater RHA than left-ear input/attention and tones. No known characteristic of these subjects accounts for this finding; JL, CB, and JM are all female and both personally and familial right-handed; DW is a personally left-handed, familial right-handed male. Note also that these "ipsilateral" qEEG patterns are not always in agreement with the behavioral EA results shown in Fig. 2: CB's behavioral EAs are REA for syllables and LEA for tones, yet her qEEG patterns show greater RHA with right-ear attention to both types of dichotic presentation.

Given this puzzling result, however, the internal consistency of the asymmetry patterns is quite good: right ear vs. left ear input/attention, and syllables vs. tones tend to show opposite asymmetries, and the interaction between ear and type of sound is similar to that seen for JL: syllables tend to evoke asymmetries in the same direction as right-ear input/attention, and tones evoke asymmetries in the same direction as left-ear input/attention.

Fig. 6 presents qEEG results for the motor activation conditions for all 4 subjects, in terms of beta power asymmery comparing electrode locations F7/F8. Note that JL shows a clear contralateral activation pattern, while the other three are consistent in their "ipsilateral" pattern. JL and CB showed resemblances between bilateral hand movement and movement of one of the other hands (R for JL, L for CB). Failure of the other two Ss to show such a match may be due to the high levels of artifact present throughout their records during these conditions, which were tested late in each session. In the future, we plan to test the somewhat fatiguing hand-flexion conditions first, while the subjects are fresh.

Fig. 6 qEEG asymmetries: Motor control
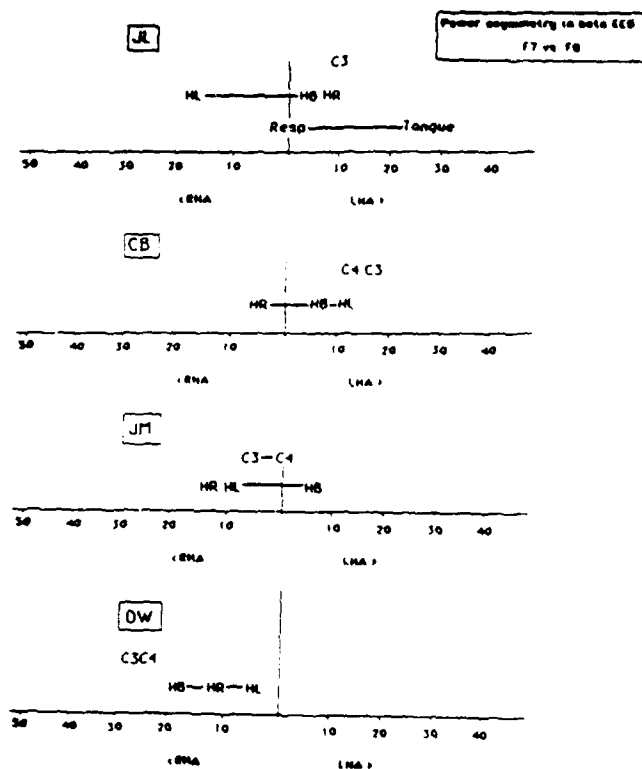
## CONCLUSIONS

Although a number of questions are generated by these data, we believe that as preliminary findings, the results are encouraging regarding the potential usefulness of qEEG as a tool for studying cerebral responses to fairly simple stimulus and task combinations, and indicate that "cognitive" processes are not the only phenomena that might be usefully studied using qEEG.

Comparisons between the EAs tested behaviorally and "hemisphere advantages" (HAs) calculated for the qEEG results, for each subject, are shown in Fig. 7. An example of the procedure used to calculate the qEEG HAs is given below the figure.

Fig. 7. Comparisons between behaviorally-determined ear advantages (lower abscissa: non-italics) and qEEG hemisphere advantages (upper abscissa: italics)
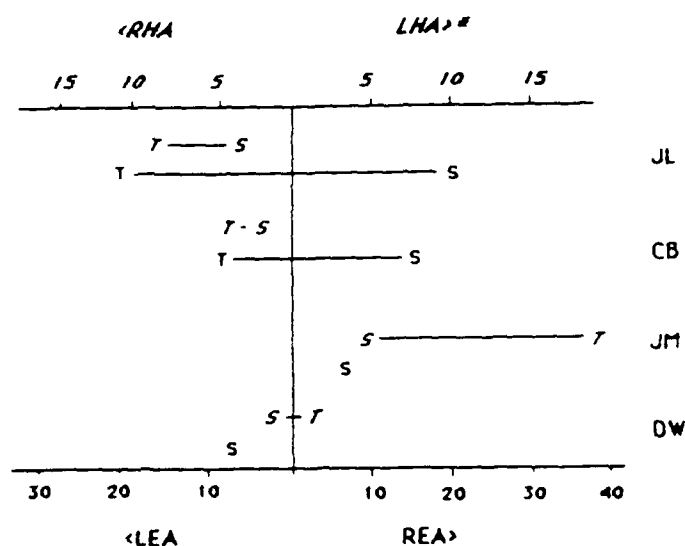


*calculated as shown in example below for JL

The degree of individual differences seen in all of these results suggests that more subjects need to be tested, particularly if we are to understand the significance of the "ipsilateral" pattern of activation shown by 3 of the 4 subjects. Future designs will also require all subjects to complete behavioral testing on all sounds before testing with qEEG

We expect that some of these puzzles will be resolved as future subjects undergo in-depth testing in our Coordinated Noninvasive Studies (CNS) Project. In this Project, subjects will first be tested behaviorally to establish each individual's ear advantages or

3 types of sounds, and then will be tested on a variety of noninvasive devices in order to observe anatomical and physiological brain asymmetries (Fig. 8: a color figure not included in this MS). Tests will include: Magnetic Resonance Imaging (MRI), Evoked Potentials (EPs—specifically, Auditory Brainstem Responses ABRs), qEEG, Magnetoencephalography (MEG), and Positron Emission Tomography (PET). Procedures will be based on our previous work with some of these devices (EPs: Lauter & Loomis 1986; In press; PET: Lauter et al 1985; 1988). During testing with each physiological device, subjects will be stimulated on separate test runs with each of the 3 sound sets. Patterns of asymmetries in measurements with the different noninvasive devices will be compared with each other, and with the behavioral asymmetries shown by the same subject (Fig. 9: a color figure not in this MS).

It is expected that the "view" of the brain available with each of the approaches will be most interpretable when considered in the context of the results on all the devices. The immediate goal of the CNS Project is to determine the degree of match between patterns of asymmetry tested behaviorally and patterns of asymmetry with regard to the same stimuli when tested using physiological methods. The ultimate goal of the Project is to take the first steps toward articulating a bridge between brain and behavior based on the new noninvasive methods, demonstrating the value of these new approaches for studying the brain by illustrating at least one way in which they may serve as the tools in a "new neuroscience," based on noninvasive methods and focused on study of the human brain.

REFERENCES

Gevins AS & RE Schaffer (1980) A critical review of electroencephalographic (EEG) correlates of higher cortical functions. CRC Crit Rev Bioeng 4: 113-164.

Gevins AS (1984) Analysis of the electromagnetic signals of the human brain: Milestones, obstacles, and goals. IEEE Trans Biomed Eng 31: 833-50.

Glass A (1984) Round table on hemispheric specialization and lateral asymmetries in the EEG: Introduction. Biol Psychol 19: 151-157.

Lauter JL (1982) Dichotic identification of complex sounds: Absolute and relative ear advantages. J Acoust Soc Amer 71: 701-707.

Lauter JL (1983) Stimulus characteristics and relative ear advantages: A new look at old data. J Acoust Soc Amer 74: 1-17.

Lauter JL (1984) Contralateral interference and ear advantages for identification of three-element patterns. Brain & Cognition 3: 259-280.

Lauter JL & RL Loomis (1986) Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability. I. Absolute latencies of brainstem vertex-positive peaks. Scand Audiol 15: 167-172.

Lauter JL & RL Loomis (1988) Individual differences in auditory electric responses: Comparisons of between-subject and within-subject variability. II. Amplitude of brainstem vertex-positive peaks. Scand Audiol 17: 88-92.

Lauter JL, P Herscovitch, C Formby and ME Raichle (1985) Tonotopic organization in human auditory cortex revealed by positron emission tomography. Hearing Res 20: 199-205.

Lauter JL, P Herscovitch, and ME Raichle (1988) Human auditory physiology studied with positron emission tomography. In J Syka and RB Masterton (Eds), Auditory pathway. NY: Plenum; pp. 313-317.

# INDIVIDUAL DIFFERENCES IN AUDITORY ELECTRIC RESPONSES: COMPARISONS OF BETWEEN-SUBJECT AND WITHIN-SUBJECT VARIABILITY

## II. *Amplitude of Brainstem Vertex-positive Peaks*

Judith L. Lauter[1] and Robert L. Loomis[2]

*From the [1]Central Institute for the Deaf, and the [2]Department of Otolaryngology,
Washington University School of Medicine, St. Louis, MO 63110, USA*

## ABSTRACT

*Individual differences in auditory electric response: comparison of between-subject and within-subject variability. II. Amplitude of brainstem vertex-positive peaks. Lauter, J. L. and Loomis, R. L. (Central Institute for the Deaf and Department of Otolaryngology, Washington University School of Medicine, St. Louis, MO., USA).*

*Scand Audiol 1988, 17 (87–92).*

Recently, we (Lauter & Loomis, 1986) reported variability measures of the latency of five vertex-positive auditory brainstem response (ABR) peaks collected under a repeated-measures experimental design. Seven subjects were tested, each on eight separate sessions, for brainstem auditory evoked response to monaural right, monaural left, and binaural stimulus presentation. This paper presents variability measures for amplitudes of the same series of responses. Three types of variability measurement were made: 1) amplitude of each peak of the response to monaural right, monaural left, and binaural stimulation; 2) amplitude difference for each peak comparing binaural with right, and binaural with left; and 3) amplitude difference comparing binaural with the sum of the amplitudes of the two monaural responses. As in the previous report, between-subject variability and within-subject variability were expressed using a ratio of mean divided by standard deviation (this is the reciprocal of Pearson's Coefficient of Variation, and will here be referred to as the Coefficient of Stability, or Cs). For all amplitude comparisons, Cs profiles indicate that: 1) within-subject stability (i.e., consistency) is significantly greater than between-subject stability, 2) both within- and between-subject stability measures are sensitive to both peak and ear of presentation, and 3) stability profiles for individual subjects show individual differences and similarities, and are replicable over time. The variability measure also provides evidence of an ear asymmetry at peak III which has been noted in other ABR studies.

## INTRODUCTION

Recent interest in the variability of auditory evoked potentials (AEPs) has been limited primarily to between-subject (BS) comparisons (cf. review in Lauter & Loomis, 1986). Our interest in the widely varying degree of individual differences to be observed in behavioral measures of sensory performance (e.g., Lauter, 1982, 1983) led us to examine the degree of within-subject (WS) variation in evoked potentials. It has long been observed that: 1) AEP waveforms for different subjects, even in ABRs, are not alike; and 2) at the same time, responses shown by individuals are quite replicable over time.

In our earlier report (Lauter & Loomis, 1986), we described latency data for a repeated-measures series of auditory brainstem response recordings, for 7 subjects each tested weekly for a total of 8 weeks. In that report, we presented both absolute and variability measures of latency, with a ratio of latency mean and standard deviation used to express latency variability of each of the five ABR peaks. Results indicated that, while absolute values showed the expected increase in latency with peak, but no differences due either to ear of presentation or to group comparisons, the variability measure proved to be sensitive to both peak and ear, as well as to BS vs WS comparisons. We also noted in that report that these ABR latency-variability 'profiles' could distinguish between subjects, and that the profiles could be replicated over time. For the current report, data from the original series were subjected to analysis for amplitude variability for each of the five peaks.

## MATERIALS AND METHODS

Auditory brainstem responses were recorded from 7 normal young adult subjects, 4 females and 3 males. Each subject was tested in eight separate weekly sessions. All sessions for each subject were scheduled for the same time of day on the same day of the week. Each subject was screened for normal hearing threshold for pure tones (better than 20 dB HL) preceding each test session; on one occasion, a subject's threshold was elevated due to a mild middle-ear infe

## ABR mean amplitude
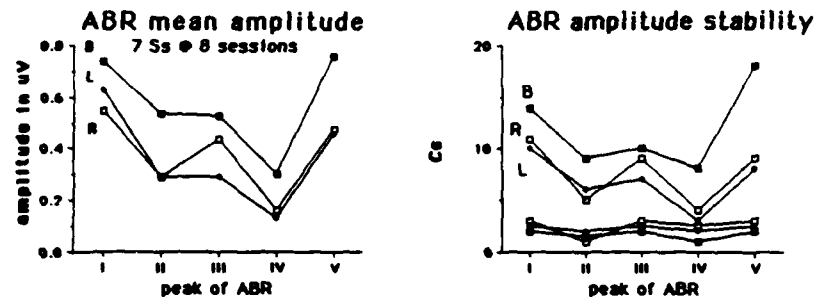
## ABR amplitude stability



*Fig. 1.* (*A*) Actual amplitude in μV for five peaks of ABR, to right-ear, left-ear, and binaural clicks (each point represents the average of 56 values: 7 Ss×8 sessions per subject). (*B*) Amplitude stability for five ABR peaks, comparing within-subject and between-subject calculations. B = binaural; R = right; L = left.

tion, and the session was rescheduled for the following week.

Each session included monaural right, monaural left, and binaural stimulation. Silver disk recording electrodes (9 mm) were placed at Cz, A1, and A2. A ground was placed at Fpz. For monaural presentations, Cz was referenced to the ipsilateral ear; for binaural, Cz was referenced to linked earlobes. Stimuli were 100 μs condensation clicks, presented at 80 dB nHL through Telex 1470 earphones with MX-41/AR cushions. ABRs were processed using a Nicolet CA-1000 system sampling once every 20 μs; records were stored on magnetic disk and analyzed off-line. Subjects reclined with eyes closed. Stimuli were presented at a rate of 11.1 clicks per second; 2 000 responses were averaged using a time window of 10 ms post-stimulus onset, and a filter setting of 150 to 3 000 Hz (−3 dB) with a 6 dB/octave roll-off. The artifact rejection criterion was 20 μV peak-to-peak.

Amplitude for each of the five peaks of the ABR waveform was defined as the range in μV between the amplitude of each vertex-positive peak and the amplitude of the following vertex-negative valley. This value was obtained for each ABR peak for each presentation mode for each subject in each session. In addition, measures were derived from these original data, for: 1) the difference between the amplitude for binaural versus right-ear stimulation (B–R), and the difference between the amplitude for binaural versus left-ear stimulation (B–L), for each peak; and 2) the difference between the amplitude for binaural stimulation versus the summed amplitudes for right-ear and left-ear stimulation (B–[R+L]), for each peak. Note that these measures were derived peak by peak, and thus do not represent the same operation as the addition and subtraction of whole waveforms as reported by Berlin and others (e.g., Berlin et al., 1984).

Amplitude values from each of these sets of data were then averaged across subjects across sessions to obtain an overall between-subject (BS) mean and standard deviation for the amplitude of each peak under all three comparisons. The amplitude values were also averaged within each subject across sessions to obtain a within-subject (WS) mean and standard deviation of the amplitude of each peak for each subject. To compare the relative variability of peak amplitudes, the ratio of mean divided by standard deviation (the reciprocal of Pearson's Coefficient of Variation), which

we will refer to as the Coefficient of Stability (Cs), was calculated for the different subject and session combinations.

## RESULTS

### Amplitude comparisons of binaural, right-ear, and left-ear conditions

Absolute amplitude, together with amplitude stability comparisons for both BS and WS calculations, are shown in Fig. 1. In panel A on the left, absolute values are plotted as a function of ABR peak. An analysis of variance for these data indicated a significant interaction between peak and ear of stimulation ($F=9.3$: $p<0.01$), with significant main effects for both peak ($F=33.11$: $p<0.01$) and ear ($F=240$: $p<0.01$).

In panel B on the right are shown variability values in terms of the coefficient of stability (Cs), comparing between-subject calculations (lower three curves) with within-subject calculations (upper three curves). There is a significant difference between the two groups of values ($F=6.77$: $p<0.05$), and there are significant interaction and main effects due to peak and ear within each set of comparisons: 1) for between-subjects, peak × ear interaction ($F=3.93$: $p<0.01$), peak main effect ($F=26.9$: $p<0.01$), ear main effect ($F=11.04$: $p<0.01$); 2) for within-subjects, there are main effects for both peak ($F=6.7$: $p<0.01$) and ear ($F=7.7$: $p<0.01$).

### Amplitude differences: binaural minus either monaural

The second amplitude comparison done for these data was the calculation of peak-by-peak differences between peak amplitude for binaural and peak amplitude for each monaural condition. Fig. 2 compares
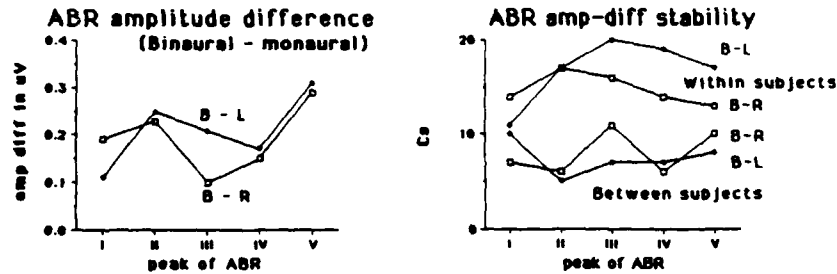
## ABR amplitude difference
(Binaural - monaural)

## ABR amp-diff stability

Fig. 2. (*A*) Actual amplitude difference in µV, comparing response to binaural vs. each monaural condition (averages of 56 values: 7 Ss×8 sessions per subject). (*B*) Amplitude

stability for the difference in µV between responses to binaural vs. each monaural condition, comparing within-subject and between-subject calculations.

absolute values (left panel) with Cs values (right panel). For absolute values, the interaction between comparison (B–R vs. B–L) and peak is significant (*F*=23.5: *p*<0.01), as are main effects for both comparison (*F*=8.6: *p*<0.01), and peak (*F*=92.86: *p*<0.01).

There is a single significant effect in the variability values shown in the right panel: the WS comparisons are more stable than the BS comparisons (*F*=72.65: *p*<0.01).

### Amplitude differences: binaural minus sum of monaurals

Fig. 3 shows the results of calculating the differences between the amplitude of each peak for the binaural condition versus the sum of that peak's amplitude for the two monaural conditions (i.e., B−[R+L]). The absolute values (on the left) show a significant effect of peak (*F*=70.35: *p*<0.01), while the variability values show significant main effects for both peak (*F*=2.83: *p*<0.05) and group (i.e., WS vs. BS: *F*=47.94: *p*<0.01).

### Individual differences

As demonstrated in our first report (Lauter & Loomis, 1986), stability profiles calculated for individuals tend to replicate over time, or depart from replication because of *increased stability* of the response at one or more peaks. Fig. 4 presents a selection of Cs curves from different individuals for a variety of amplitude measures. These illustrate characteristics of individual Cs profiles which we have observed for stability measures of both latency and amplitude: 1) good replicability from the first four weeks to the second four weeks of sessions: 2) pattern enhancement over time, which often takes the form of 3) patterns for one individual coming to resemble those of another subject.

Each curve in Fig. 4 represents the stability of the indicated measure (e.g. left-ear amplitude) for each subject calculated over four test sessions. Each pair of curves shown here represents good replication (i.e., there are no significant differences when tested with paired *t*-tests). These curves also demonstrate two other characteristics of Cs profiles: 1

## ABR amplitude difference
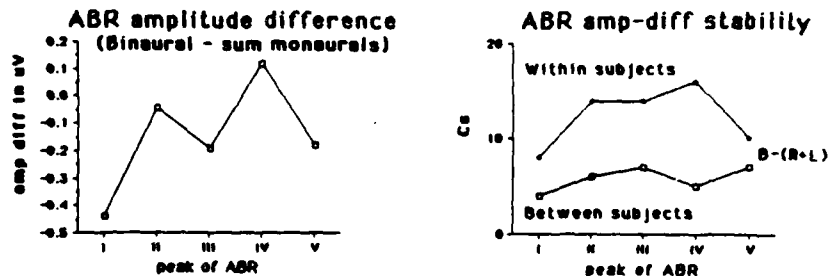(Binaural - sum monaurals)

## ABR amp-diff stability

Fig. 3. (*A*) Actual amplitude difference in µV, comparing response to binaural vs. sum of monaural conditions (averages of 56 values: 7 Ss×8 sessions per subject). (*B*) Amplitude stability for the difference in µV between re-

sponses to binaural vs. sum of responses to monaural conditions, comparing within-subject and between-subject calculations.

## Replication of ABR amplitude stability profiles
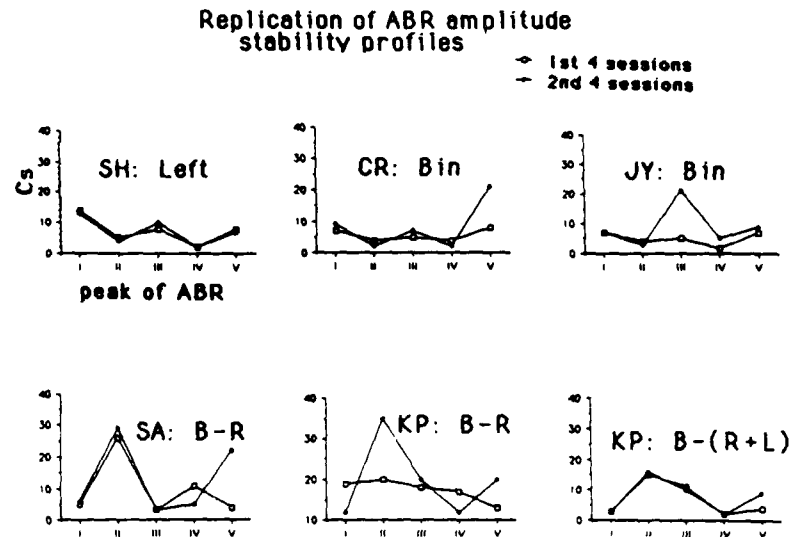
→ 1st 4 sessions
→ 2nd 4 sessions



*Fig. 4.* Replication of ABR amplitude stability profiles for 5 subjects and four measures: response to left-ear clicks, response to binaural clicks, difference between response to binaural vs. right-ear clicks, and difference between response to binaural vs. sum of responses to monaural clicks.

changes in profile shape are most often due to increased stability at one or more peaks (cf. CR Binaural peak V, JY Binaural peak III, SA Binaural – Right peak V, and KP Binaural – Right peaks II and V): and 2) one subject's profile may require the second month to assume the shape that another individual's profile showed even in the first month (compare SA's Binaural – Right profile, well-replicated over 2 months, with KP's Binaural – Right, which is flat for month one, but during the second test month, acquires a similar shape to the SA profile.

### Amplitude asymmetries

Several reports have documented asymmetries in the ABR. Berlin et al. (1984) and others have used waveform addition and subtraction methods to demonstrate an ABR asymmetry, the Binaural Interaction Component, which in some subjects comprises a difference component occurring at approximately 6 ms post-stimulus onset. Levine & McGaffigan (1983) have described an asymmetry in peak amplitude, distinguishing the absolute amplitude of responses to left- vs. right-ear stimulation, which was especially prominent at ABR peak III.

Examination of our data for both the absolute values and the stability of ABR peak amplitudes (cf. Figs. 1, 2, and 3) suggests that our findings also are sensitive to response asymmetries, perhaps directly

related to the feature described by Levine & McGaffigan (1983). For example, in Fig. 1, panel A, the absolute amplitude of the right-ear response at peak III is clearly dominant over that of the left-ear response: the direction and magnitude of this difference for our group of subjects (all right-handed), amounting to a right-ear preference of approximately 0.11 μV, compares well with that reported by Levine & McGaffigan (1983) for their subset of right-handed subjects, whose data showed a group average of about 0.08 μV difference between right and left amplitude.[1] Panel B shows a similar advantage at peak III for right-ear stimulation, however, in this case, in terms of greater stability of the amplitude of the right-ear versus the left-ear response. Note that for both the absolute and the stability measure, the 'advantage' of the right-ear response amounts to its approaching the corresponding value for the binaural response—whether in absolute amplitude (panel A) or in stability (panel B).

---

[1] Levine & McGaffigan (1983) measured amplitude as baseline-to-positive peak. The right–left difference in this amplitude measure at peak III, read from Fig. 1 B of their paper, and averaged over their right-handed subjects, was about 0.04 μV. For comparison with our data, we have doubled this value as a rough approximation to our peak-to-valley definition of amplitude, to yield a right–left difference in peak III amplitude for their right-handed subjects of about 0.08 μV.

This similarity between binaural and right-ear response at peak III for this group of subjects is further corroborated by the comparisons shown in Fig. 2. In panel A, the difference between actual amplitudes of binaural versus monaural responses shows a greater separation between binaural and left-ear responses at peak III than between binaural and right-ear responses: the gap between binaural and left-ear peak III amplitudes for these subjects is almost twice that separating binaural and right-ear: 21 μV for (B–L) versus 10 μV for (B–R). Related patterns are seen in the stability data in panel B. For the between-subject profiles, the peak III difference between (B–R) and (B–L) is clear, with the B–R difference the more stable of the two. Differences in stability of the two binaural/monaural comparisons for within-subject calculations are more complex, with stability asymmetry this time favoring the B–L difference at peaks IV and V as well as III.

In Fig. 3, the absolute values in panel A point to the importance of right vs. left differences at peak III, where the B−[R+L] values show a dramatic drop. This is related to the difference between the absolute-value curves in Fig. 2, panel A, and reflects the similarity of right-ear and binaural amplitude at peak III in these subjects. Stability of this calculation (panel B) shows no differences for peak III.

## DISCUSSION

In general, these results on ABR amplitude stability are in keeping with the findings on the stability of ABR peak latencies reported in Lauter & Loomis (1986). Both analyses illustrate the dramatic increase in information regarding the ABR waveform to be gained by: 1) conducting evoked-potential testing under a repeated-measures design, and then 2) studying waveform parameters such as latency and amplitude in terms of both absolute and stability measures rather than in terms of absolute measures alone.

For both the ABR latency and amplitude data, within-subject stability (i.e., consistency) is generally greater than between-subject stability. In addition, within-subject group stability profiles as well as individual profiles reveal detailed interactions between peak and ear of stimulation contributing ⁺o distinctions in the degree of waveform-parameter stability at different waveform peaks. Further analysis of individual differences in evoked-potential testing, perhaps considered in combination with be-

havioral results from the same subjects, may lead to an understanding of the mechanisms underlying Cs patterns. These types of comparisons should be of particular interest for studying correlations between asymmetries of response demonstrated electrophysiologically, with those that can be observed with behavioral tests such as dichotic listening (e.g., Lauter, 1982; 1983).

Consideration of individual patterns in terms of group characteristics of these stability profiles may provide additional insights. As described above, resemblances in Cs patterns can be observed between subjects; in our data, such resemblances were sometimes visible in the first month of testing, and in other cases, were not apparent until the second test month. Extended testing of the same subjects, perhaps comparing EP waveforms collected according to different time schedules (four waveforms per day versus four per week versus four per month, etc.), should clarify the patterns of individual response, including the degree to which individuals differ in their stability profiles, the groupings that are possible based on profile types, and the time course required for each subject to reveal his/her characteristic profile for a given ABR parameter.

We expect that further consideration of the patterns of EP response stability, including extended time of testing, examination of changes in stability patterns with age, and comparisons between the results of electrophysiological and behavioral testing in the same subjects, will provide new insights into the organization of the human auditory system, as well as guidelines for clinical applications of repeated-measures evoked-potential testing.

## REFERENCES

Berlin, C. I., Hood, L. J. & Allen, P. 1984. Asymmetries in evoked potentials. In *Hearing science: recent advances* (ed. C. I. Berlin), pp. 461–477. College-Hill Press, San Diego.

92     *J. L. Lauter and R. L. Loomis*

Lauter, J. L. 1982. Dichotic identification of complex sounds: Absolute and relative ear advantages. *J Acoust Soc Am 71*, 701.

Lauter, J. L. 1983. Stimulus characteristics and relative ear advantages: A new look at old data. *J Acoust Soc Am 74*, 1.

Lauter, J. L. & Loomis, R. L. 1986. Individual differences in auditory electric responses: comparisons of between-subject and within-subject variability. I. Absolute latencies of brainstem vertex-positive peaks. *Scand Audiol 15*, 167.

Levine, R. B. & McGaffigan, P. M. 1983. Right–left asymmetries in the human brain stem: auditory evoked potentials. *Electroenceph Clin Neurophysiol 55*, 532.

Address for offprints:

Judith L. Lauter
Dept. of Speech and Hearing Sciences
University of Arizona
Tucson AZ 85721, U.S.A.

HUMAN AUDITORY PHYSIOLOGY STUDIED WITH POSITRON EMISSION TOMOGRAPHY

Judith L. Lauter*, Peter Herscovitch**, Marcus E. Raichle**

*Speech and Hearing Sciences
University of Arizona
Tucson AZ 85721 USA
**Edward Mallinckrodt Institute of Radiology
Washington University School of Medicine
St. Louis, MO 63110, USA

Although the past 40 years have seen significant progress in our
understanding of the organization of sensory nervous systems in a number
of animals, access to the details of human sensory CNS structure and func-
tion has been hampered by the lack of noninvasive, high-resolution
technology. However, within the last decade, a number of new devices have
appeared that provide relatively noninvasive access to the human brain:
e.g. CT and MRI for anatomical imaging, and MEG, BEAM, and PET for topo-
graphic physiological studies.

METHODS

PET

Positron emission tomography represents a modification of tissue auto-
radiography techniques, and depends on the ability of radiation products
of positron-emitting isotopes to penetrate the human skull, and thus
become externally detectable. When a bolus of water labelled with oxygen-
fifteen is injected into a subject's arm vein, a ring of detectors surround-
ing the subject's head can generate a data array that can be used to
reconstruct an image showing a topography of greater and lesser concentra-
tions of isotope. The resulting images represent the brain as a series of
slices ranging from the top of the brain down into cerebellum. A color
scale indicates regions of greater and lesser isotope concentration, or,
if blood samples are taken during scanning to monitor actual isotope
levels, regions of greater and lesser blood flow. Using appropriate
software, such images can also be combined to produce difference images,
showing derived maps of areas undergoing greater or lesser change in blood
flow/isotope concentration from control scan to a scan taken under stimula-
tion conditions.

For these studies, positron emission tomography was performed using
a PETT VI system (Ter-Pogossian et al., 1982; Yamamoto et al., 1982).
Data are recorded simultaneously for 7 slices with a center-to-center
separation of 14.4 mm; the in-plane (i.e., transverse) reconstructed
resolution is about 12.4 mm in the center of the field of view, and slice
(axial) thickness is about 13.9 mm at the center. Each scan is 40 sec in

313

length, and is performed following the intravenous bolus injection of about 10 ml of saline containing 55-80 mCi of O-fifteen-labelled water (half life: 123 sec). Cerebral blood flow (CBF: ml/(min x 100 g)) is calculated using a PET adaptation of the Kety tissue autoradiographic technique previously described and validated in our laboratory (Herscovitch et al., 1983; Raichle et al., 1983).

For auditory studies, we have designed a sound-delivery system based on insert receivers set in plastic tubing that snaps into standard earmolds. This fits underneath the face mask (see below) , and allows not only the shielding of text stimuli from ambient noise, but also the isolation of sounds to the two ears, to distinguish monaural, binaural, and dichotic presentations. The frequency response of this system has been shaped to mimic the filter characteristics of the outer ear, so that the signal presented at the eardrum is "ecologically valid" in its acoustical makeup (see Lauter et al., 1985 for a complete description).

## Subjects

Normal young adults with no history of neurological or hearing disorders served as subjects; each was paid for his/her participation. Prior to testing, each subject received an orientation visit to the laboratory, when all procedures were explained, and a consent form was read and signed.

Subject preparation preceding each session included the percutaneous insertion of a radial arterial catheter, under local anesthesia, to permit frequent sampling of arterial blood during scans, and the insertion of an intravenous catheter in the opposite arm for isotope injection. The head was positioned with a special head holder which utilized an individual molded plastic face mask to prevent movement during the study. A laser permanently attached to the wall projected a line onto the mask that corresponded to the position of the lowest PET slice. A lateral skull radiograph with this line marked by a radiopaque wire provided a record of the subject's exact position in relation to the PET slices. The overlapping position of radiopaque markers placed in the external auditory canals (the earmold rings) confirmed that the head was not rotated about the anterior-posterior or vertical axes. After the head was in place, a transmission scan used for individual attenuation correction was performed with a ring source of activity containing germanium-68/gallium -68. During scans the room was darkened and the subject's eyes were covered with gauze pads. Ambient noise during each scan was limited to the sound of cooling fans for the electronic equipment.

## Stimuli

A variety of sounds has been used in our test series. Results to be reviewed here will focus on experiments using pure tones and synthetic syllables.

Pure tones. Pure tones were generated using a General Radio 1310A oscillator, an electronic switch and pulse generator built at Central Institute for the Deaf in St. Louis, and a Hewlett-Packard 350D attenuator. Tones were monitored using a Monsanto 113A counter, Telequipment S54A oscilloscope and Hewlett-Packard 400GL voltmeter.

Tones of 500 Hz and 4 kHz were used for testing. Tone were pulsed with a duty cycle of 50%, approximately 500 msec on/off, with a rise/fall time of 50 msec. The subject's threshold for each frequency tested was determined just prior to scanning for that frequency. All tones were presented at 50 dB SL, monaurally to the right ear. For each experimental

314

scan, the sound was turned on approximately 1 min prior to isotope injection, and was presented throughout the scan; thus total presentation time was approximately 2 min.

Synthetic syllables. A tape recording of a set of synthetic nonsense stop-consonant-vowel (stop CV) syllables used in our dichotic listening experiments (e.g., Lauter, 1982) was presented to subjects via a Nagra tape recorder. In preparing this recording, the original 250-msec version of each syllable was edited to leave only the first 50 msec, including acoustical information regarding both consonant and vowel. The tape recording consisted of a constant cycling of the syllable string (ba-da-ga-pa-ta-ka-ba-da-ga...etc.). The rate of syllable repetition, overall level of the recording, and ear of presentation were manipulated in separate experiments (see below). As with the tones, the subject's threshold for the tape recording was obtained just prior to the text scan.

## Anatomical Localization

In order to determined where in the three-dimensional data complex to look for auditory responses, we used an anatomical localization scheme developed in our laboratory (cf Fox et al., 1984) that is independent of the appearance of the CBF images. This method yields both slice number and transverse-plane coordinates for a predicted region of interest (ROI) selected from a standard stereotaxic atlas of the human brain (Talairach et al., 1967).

We identified two ROIs for the scans involving pure tones and syllables: primary auditory cortex, and a region surrounding the angular gyrus, often designated as "language cortex." Tomographic images from each subject were then used to create "percent-difference images" (Fox and Raichle, 1984), comparing control and experimental conditions. These images are based on blood-flow values normalized to control for global changes in blood flow occurring between scans, and to highlight areas of maximum change from control to stimulated condition that occur independent of any global changes in CBF.

## RESULTS

Pure tones. Examination of activity changes within the estimated region of primary auditory cortex for each hemisphere of each subject revealed systematic shifts of the area of maximum change from condition to condition. In each subject, maximum change always occurred in the left-hemisphere AI region (i.e., contralateral to stimulation). Also, for each subject, the contralateral region of greatest activity change during stimulation with the 500 Hz tone was more lateral and anterior, and the region that responded best to the 4 kHz tone was more medial and posterior. The orientation of these regions for the five subjects tested in six sessions agree well with those reported for tonotopic responses in monkey auditory cortex using electrophysiological methods (e.g., Brugge and Merzenich, 1973). (See Lauter et al., 1985 for a complete description of these results.)

Synthetic syllables. Results are available to date for single-subject examples of the effects of manipulating rate, level, and ear of presentation of the recorded syllables. Clear qualitative changes were observed in the rCBF images in response to the syllables, occurring in the angular-gyrus "language cortex" region previously defined for each subject. Analysis of the quantitative changes in rCBF as a function of the dimensional manipulations indicate that as rate and level are increased, there is a corresponding increase in rCBF; as ear of presentation is changed, related shifts in activation seem to reflect the predominance of
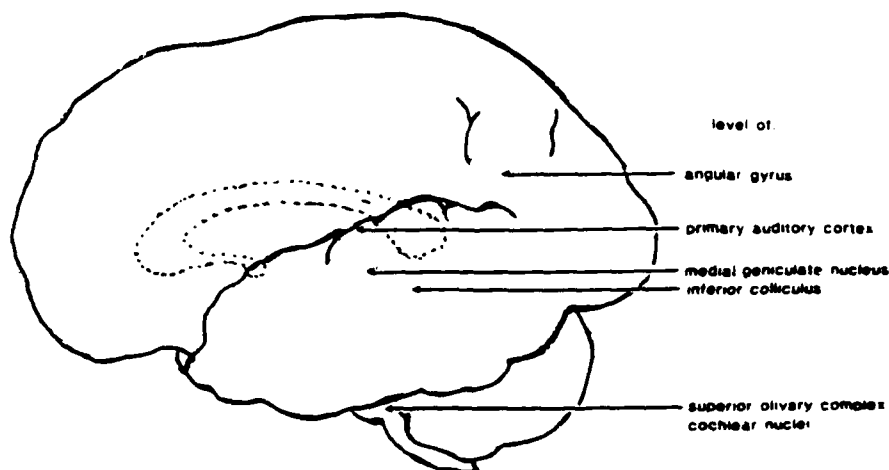
Fig. 1. Schematic of nuclear levels within the human auditory CNS.
in these results, but since they are based on data from single subjects,
the suggestions cannot be taken as conclusive.

contralateral response. There are suggestions of asymmetrical responses

Multi-level activation of auditory nuclei. The human auditory CNS is
advantageously arranged for study with a tomographic device such as the
PETT VI (Fig. 1). This is in contrast, for example, with the visual
system, which lies essentially within the dimensions of a single PET
slice. As a result, it might be possible to view, in a single 40-sec scan,
responses in more than one auditory center to a particular stimulus.

Fig. 2 presents a series of difference images taken in a single scan
of one of our test subjects, representing comparisons between a control
scan and a stimulation scan in which synthetic syllables were presented
binaurally at a rate of 20 per second at a level of 50 dB SL. The regions
represented in the two most rostral slices (Panel A, upper two images)
contain no known auditory centers. The slice shown in the lower left of
Panel A, however, at the level of the angular gyrus for this subject,
shows clear bilateral activation, and there is an apparent asymmetry of
substantial proportions--a 16% "left hemisphere advantage" in terms of
rCBF change. Further analysis will be required to determine whether this
difference is statistically significant. The lower-right slice is at the
level of primary auditory cortex; note again bilateral activation, more
symmetrical at this level. The top left slice of Panel B represents the
level of the thalamus: the striking bilateral, symmetrical activation
seen here may be interpretable as response in posterior thalamus, perhaps
representing a combination of MGN and pulvinar. The top right slice is at
the level of the midbrain; the midline activation could indicate inferior
colliculus response, with separation of the two halves of the IC beyond
the resolution of PETT VI. The last slice may be through the cerebellum;
significance of the small response seen here is unknown.

DISCUSSION

This library of PET auditory activation studies in normal human brains
will be used to answer a variety of additional questions. However, results
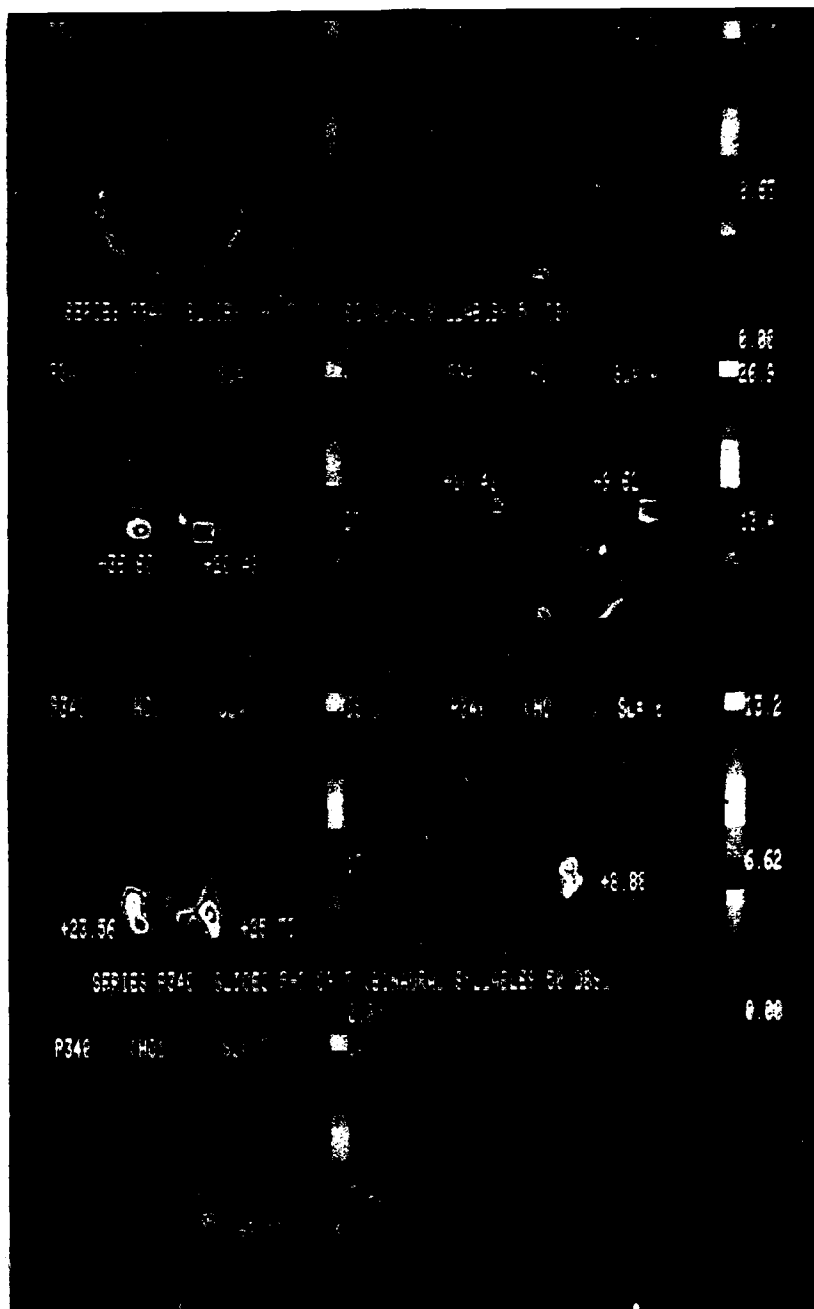
316

Fig. 2.  Simultaneous multi-level activation of human auditory CNS. Series
        of 7 slices (Panel A, SL 1 most rostral, Panel B, SL 7 most
        caudal) from a single 40-sec scan on one subject. Slices are shown
        as "difference images", comparing control with stimulation by
        synthetic syllables presented binaurally at a rate of 20 per
        second at a level of 50 dB SL. (See text for details.)

to date already suggest that positron emission tomography holds enormous potential as a revolutionary tool for the study of normal human sensory physiology. In the auditory system, responses to both simple and complex sounds can be observed, at a number of auditory CNS levels, with brain activity integrated over as little as 40 sec. The new generation of PET machines (e.g., "SuperPet") will provide improved spatial resolution and much better temporal resolution, sufficient for "evoked rCBF response" studies.

We believe that our results with auditory stimulation, combined with parallel findings in other modalities, point to the possibility of a new physiology, pursued via noninvasive techniques, and designed to emphasize human nervous systems and the complex interactions between stimulus, presentation, and subject variables that are the hallmark of everyday behavior.

## REFERENCES

Brugge, J. F. and Merzenich, M. M., 1973, Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation, J. Neurophysiol., 36:1138-1158.

Fox, P. T. and Raichle, M. E., 1984, Stimulus rate dependence of regional cerebral blood flow in human striate cortex, demonstrated by positron emission tomography, J. Neurophysiol., 51:1109-1120.

Fox, P. T., Perlmutter, J. S., and Raichle, M. E., 1984, A stereotactic method of anatomical localization for positron emission tomography, J. Comp. Assist. Tomogr., 9:141-153.

Herscovitch, P., Markham, J., and Raichle, M. E., 1983, Brain blood flow measured with intravenous H2-O15. I. Theory and error analysis, J. Nucl. Med., 24:782-789.

Lauter, J. L., 1982, Dichotic identification of complex sounds:Absolute and relative ear advantages, J. Acoust. Soc. Amer., 71:701-707.

Lauter, J. L, Herscovitch, P., Formby, C., and Raichle, M. E., 1985, Tonotopic organization in human auditory cortex revealed by positron emission tomography, Hearing Res., 20:199-205.

Reichle, M. E., Martin, W. R. W., Herscovitch, P., Mintun, M. A., and Markham, J., 1983, Brain blood flow measured with intravenous H2-O15. II. Implementation and validation, J. Nucl. Med., 24:790-798.

Talairach, J., Szikla, O., and Tournous, P., 1967, Atlas d'Anatomie stereotaxique du Telencepahle, Masson, Paris.

Ter-Pogossian, M. M., Ficke, D. C., Hood, J. T., Yamamoto, M. and Mullani, N. A., 1982, PETT VI: Positron emission tomography utilizing cesium fluoride scintillation detectors, J. Computer Assist. Tomogr., 6:125-133.

Yamamoto, M., Ficke, D. C., Ter-Pogossian, M. M., 1982, Performance study of PETT VI, a positron computed tomograph with 288 cesium fluoride detectors, IEEE Trans. Nucl. Sci. NS, 29:529-533.